

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 16.833 Seconds
(without alignments)
1277.688 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225

Sequence: 1 PLYSYTDCGNTVTSSDGL.....NNQTQFLFLEHLLTANTTKG 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	100.0	359	2 O43656	O43656 homo sapien
2	225	100.0	364	2 Q96D42	Q96D42 homo sapien
3	178	79.1	451	2 Q95144	Q95144 cercopithec
4	178	79.1	460	2 O18984	O18984 cercopithec
5	178	79.1	469	2 Q7J348	Q7J348 cercopithec
6	178	79.1	473	2 Q7J347	Q7J347 cercopithec
7	178	79.1	474	2 O46597	O46597 cercopithec
8	178	79.1	478	2 O46598	O46598 cercopithec
9	62.5	27.8	639	1 YHR7 YEAST	P38825 saccharomyc
10	62	27.6	842	2 Q9U0J9	Q9U0J9 plasmodium
11	61.5	27.3	1269	2 O97Z06	O97Z06 sulfolobus
12	61	27.1	1255	2 O36287	O36287 venezuelan
13	60.5	26.9	3056	2 Q7USQ0	Q7USQ0 rhodospirell
14	60	26.7	602	1 NRGI CHICK	Q05199 gallus gall
15	59	26.2	142	2 Q6KPY4	Q6KPY4 rickettsia
16	59	26.2	549	2 O53969	O53969 shigella dy
17	59	26.2	559	2 Q36298	Q36298 venezuelan
18	59	26.2	626	2 Q6KHE7	Q6KHE7 mycoplasma
19	59	26.2	1254	1 P0LS BEWE	P36330 venezuelan
20	59	26.2	1254	2 Q77VY8	Q77VY8 venezuelan
21	59	26.2	1259	2 Q9YKCS	Q9YKCS venezuelan
22	59	26.2	2951	2 Q8THC9	Q8THC9 methanosarc
23	58.5	26.0	931	2 Q6WP59	Q6WP59 dengue viru
24	58.5	26.0	931	2 Q6WP60	Q6WP60 dengue viru
25	58.5	26.0	931	2 Q6WP63	Q6WP63 dengue viru
26	58.5	26.0	931	2 Q6WP68	Q6WP68 dengue viru
27	58.5	26.0	931	2 Q6WP69	Q6WP69 dengue viru
28	58.5	26.0	931	2 Q6WP70	Q6WP70 dengue viru
29	58.5	26.0	1258	2 Q9WC27	Q9WC27 venezuelan
30	58.5	26.0	3392	2 Q8JQD9	Q8JQD9 dengue viru
31	58.5	26.0	3392	2 Q91NH1	Q91NH1 dengue viru

RESULT 1

O43656 PRELIMINARY; PRT; 359 AA.

AC O43656; 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)

DE Hepatitis A virus cellular receptor 1.

GN Name=HAVcr-1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=98325180; PubMed=9658108;

RA Feigelstock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;

RT "The human homolog of HAVcr-1 codes for a hepatitis A virus cellular

RT receptor.";

RL J. Virol. 72:6621-6628(1998).

DR EMBL; AF043724; AAC39862.1; -.

DR GO; GO:0004872; F:receptor activity; IEA.

DR InterPro; IPR003599; IG.

DR InterPro; IPR007110; IG-like.

DR SMART; SM00409; IG_MHC.

DR PROSITE; PS50835; IG LIKE; 1.

DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

KW Receptor.

SQ SEQUENCE 359 AA; 38704 MW; C207FEC562DC62CA CRC64;

Query Match 100.0%; Score 225; DB 2; Length 359;

Best Local Similarity 100.0%; Pred. No. 2.8e-20;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLYSYTDCGNTVTSSDGLMNNNTQFLFLEHLLTANTTKG 42

Db 249 PLYSYTDCGNTVTSSDGLMNNNTQFLFLEHLLTANTTKG 290

RESULT 2

O96D42 PRELIMINARY; PRT; 364 AA.

AC O96D42; 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)

DE HAVCR1 protein.

GN Name=HAVCR1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

OX NCBI_TaxID=9606;

RN SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards D.K., Worley K.C., Hale S.G., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitling M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC013325; AAH13325.1; -;
 DR EMBL; CR457114; CAG33395.1; -;
 DR Genew; HGNC:17866; HAVCR1.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-Like.
 DR InterPro; IPR003006; Ig_MHC.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 SQ SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;
 Query Match 100.0%; Score 225; DB 2; Length 364;
 Best Local Similarity 100.0%; Pred. No. 2.8e-20; Mismatches 0; Indels 0; Gaps 0;
 Matches 42; Conservative 0;
 QY 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLTANTTKG 42
 Db PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLTANTTKG 295
 RESULT 3
 Q95144 PRELIMINARY; PRT; 451 AA.
 AC Q95144;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DE HAVCR-1 protein precursor.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=97015129; PubMed=8861957;
 RA Kaplan G., Totsuka A., Thompson P., Akatsuka T., Moritsugu Y.,
 RA Feinstein S.M.;
 RT "Identification of a surface glycoprotein on African green monkey

RT kidney cells as a receptor for hepatitis A virus.";
 RL EMBO J. 15:4282-4296(1996).
 DR EMBL; X98252; CAA66906.1; -;
 DR PIR; S71754; S71754.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-Like.
 DR InterPro; IPR003006; Ig_MHC.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Signal.
 FT SIGNAL 1 17 Potential.
 SQ SEQUENCE 451 AA; 48774 MW; 5D395C5455AA4332 CRC64;
 Query Match 79.1%; Score 178; DB 2; Length 451;
 Best Local Similarity 81.0%; Pred. No. 3.6e-14; Mismatches 6; Indels 0; Gaps 0;
 Matches 34; Conservative 2;
 QY 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLTANTTKG 42
 Db PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLTANTTKG 370
 RESULT 4
 O18984 PRELIMINARY; PRT; 460 AA.
 AC O18984;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hepatitis A virus receptor.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97368416; PubMed=9225030;
 RA Ashida M., Hamada C.;
 RT "Molecular cloning of the hepatitis A virus receptor from a simian
 RT cell line.";
 RL J. Gen. Virol. 78:1565-1569(1997).
 DR EMBL; D88585; BAA21556.1; -;
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-Like.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 460 AA; 49687 MW; 1D1A0C1D832139EA CRC64;
 Query Match 79.1%; Score 178; DB 2; Length 460;
 Best Local Similarity 81.0%; Pred. No. 3.7e-14; Mismatches 6; Indels 0; Gaps 0;
 Matches 34; Conservative 2;
 QY 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLTANTTKG 42
 Db PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLTANTTKG 379
 RESULT 5
 Q7J48 PRELIMINARY; PRT; 469 AA.
 AC Q7J48;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Hepatitis A virus cellular receptor 1 short form.
 GN Name=HAVcr-1;
 OS Cercopithecus aethiops (Green monkey) (Grivet).

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC	Cercopitheciinae; Cercopithecus.
ON	NCBI_TaxID=9534;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98285791; PubMed=9621093;
RA	Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT	"Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT	green monkey kidney cells result in antigenic variants that do not
RT	react with protective monoclonal antibody 190/4.";
RL	J. Virol. 72:6218-6222(1998).
DR	ENBL; AF043446; AAC39771.1; -.
GO	GO:0004872; F:receptor activity; IEA.
DR	InterPro; IPR003599; Ig.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003006; Ig_MHC.
DR	SMART; SM00409; IG; 1.
DR	PROSITE; PS50835; IG LIKE; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW	Receptor.
SQ	SEQUENCE 469 AA; 50561 MW; 27E386F47948F528 CRC64;
Query Match 79.1%; Score 178; DB 2; Length 469;	
Best Local Similarity 81.0%; Pred.No.3.7e-14;	
Matches 34; Conservative 2; Mismatches 6; Indels 0; Gaps 0;	
QY	1 PLSYTTDGDNDVTVESSDGLWNNTQLFLHSLLTANTTKG 42 : : : : : :
DB	347 PLSYTTDGDNDVTVESSDGLWNNTQLSPHSQPMVNTTEG 388 : : : : : :
RESULT 6	
Q7JJ47	PRELIMINARY; PRT; 473 AA.
ID	Q7JJ47
AC	Q7JJ47; 2004 (TrEMBLrel. 27, Created)
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Hepatitis A virus cellular receptor 1 short form.
GN	Name=HAVcr-1;
OS	Cercopithecus aethiops (Green monkey) (Grivet).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC	Cercopitheciinae; Cercopithecus.
ON	NCBI_TaxID=9534;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=98285791; PubMed=9621093;
RA	Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT	"Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT	green monkey kidney cells result in antigenic variants that do not
RT	react with protective monoclonal antibody 190/4.";
RL	J. Virol. 72:6218-6222(1998).
DR	ENBL; AF043448; AAC39773.1; -.
GO	GO:0004872; F:receptor activity; IEA.
DR	InterPro; IPR003599; Ig.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR003006; Ig_MHC.
DR	SMART; SM00409; IG; 1.
DR	PROSITE; PS50835; IG LIKE; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW	Receptor.
SQ	SEQUENCE 473 AA; 50973 MW; CD15EF5BE79C8013 CRC64;
Query Match 79.1%; Score 178; DB 2; Length 473;	
Best Local Similarity 81.0%; Pred.No.3.8e-14;	
Matches 34; Conservative 2; Mismatches 6; Indels 0; Gaps 0;	
QY	1 PLSYTTDGDNDVTVESSDGLWNNTQLFLHSLLTANTTKG 42 : : : : : :
DB	351 PLSYTTDGDNDVTVESSDGLWNNTQLSPHSQPMVNTTEG 392 : : : : : :

```
KW Receptor.
SQ SEQUENCE 478 AA; 51501 MW; 3A919655C752FF60 CRC64;

Query Match 79.1%; Score 178; DB 2; Length 478;
Best Local Similarity 81.0%; Pred. No. 3.8e-14;
Matches 34; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVESSDGLWNNNTQQLFLEHSLTLTANTTKG 42
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 356 PLYSYTTDGNVTVESSDGLWNNNTQQLSPHSPQWNTTEG 397

RESULT 9
YHR7 YEAST STANDARD; PRT; 639 AA.
AC P36825;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DE Hypothetical 71.9 kDa protein in ERP5-ORC6 intergenic region.
DE Hypothetical 71.9 kDa protein in ERP5-ORC6 intergenic region.
GN OrderedLocusNames=YHR117W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hallier L.W., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevasakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RL Science 265:2077-2082(1994)
RC -!- SIMILARITY: Contains 9 TPR repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00059; AAB68868.1; -.
DR PIR; S48959; S48959.
DR HSSP; P31948; IELW.
DR GerMOnline; I39434; -.
DR SGO; S000001159; TOM71.
DR GO; GO:0005741; C:mitochondrial outer membrane; IDA.
DR GO; GO:0008565; F:protein transporter activity; IPI.
DR InterPro; IPR001440; TPR.
DR InterPro; IPR008941; TPR-like.
DR Pfam; PF00515; TPR; 8.
DR SMART; SM00028; TPR; 7.
DR PROSITE; PS00005; TPR; 7.
DR PROSITE; PS50293; TPR_REGION; 2.
KW Hypothetical protein; Repeat; TPR repeat.
FT REPEAT 127 160
FT REPEAT 162 194
FT REPEAT 162 194
FT REPEAT 345 378
FT REPEAT 379 411
FT REPEAT 412 445
FT REPEAT 447 479
FT REPEAT 480 513
FT REPEAT 530 563
FT REPEAT 565 597
SQ SEQUENCE 639 AA; 71856 MW; B4AD583B6CB466B CRC64;

Query Match 27.8%; Score 62.5; DB 1; Length 639;
Best Local Similarity 40.5%; Pred. No. 29;
Matches 17; Conservative 6; Mismatches 4; Indels 15; Gaps 2;

QY 2 LYSYTTDGNVTVESSDGLWNNNTQQLFLEHSLTLTANT 39
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 305 LYSATDEGYLVANDLLTKSTD-----MYHSLLSANT 335

RESULT 10
Q9U0J9 PRELIMINARY; PRT; 842 AA.
AC Q9U0J9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Hypothetical protein PF00430C.
GN Name=PF00430C;
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]_
RP SEQUENCE FROM N.A.
RX MEDLINE=2255708; PubMed=12368867; DOI=10.1038/nature01095;
RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
RA Feltwell T., Goble A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
RA Harper D., Hauser H., Horneby T., Holroyd S., Horrocks P.,
RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
RA Rajadream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
RA Sultson J.E., Craig A., Newbold C., Barrell B.G.;
RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13.";
RL Nature 419:527-531(2002).
RN [2]_
RP SEQUENCE FROM N.A.
RA Devlin K., Pain A., Berriman B., Hall N., Bowman S., Churcher C.,
RA Harris B., Harris D., Lawson D., Quail M., Barrell B.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035475; CABG2864.1; -.
DR InterPro; IPR001862; MAC_perforin.
DR Pfam; PF01823; MACPF; 1.
KW Hypothetical protein.
SQ SEQUENCE 842 AA; 94515 MW; B6C7FA0EF9B956C7 CRC64;

Query Match 27.6%; Score 62; DB 2; Length 842;
Best Local Similarity 34.2%; Pred. No. 46;
Matches 13; Conservative 7; Mismatches 18; Indels 0; Gaps 0;

QY 4 SYTTDGNVTVESSDGLWNNNTQQLFLEHSLTLTANTTK 41
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 SSSNDNDEENEEQDDVMDNDQNDKDKIKHSFNLANESK 180

RESULT 11
Q97Z06 PRELIMINARY; PRT; 1269 AA.
AC Q97Z06;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=SSO1141;
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
```



```

OC Sulfolobus.
OX NCBI_TaxID=2287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA De Moors A., Brauso G., Fletcher C., Clausen I.G., Curtis B.A.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
DR EMBL; AF006730; AAK41392.1; -.
DR PIR; A90267; A90267.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004289; F:subtilase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001611; LRR.
DR PRINTS; PR00019; LEURICRPT_5.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Complete proteome; Protease.
SQ SEQUENCE 1269 AA; 138027 MW; 87C8F05CCF83DA84 CRC64;

Query Match 27.3%; Score 61.5; DB 2; Length 1269;
Best Local Similarity 34.0%; Pred. No. 85;
Matches 18; Conservative 7; Mismatches 11; Indels 17; Gaps 2;

OY 4 SYTDG-----NDVTE-SSDGLMNNQTLFLEHSLTANT 39
DB 1163 SPTTNGTHIVTINTQNPYDGYNLTVTAIQDGLSSSSSYLVFENGTLNLT 1215

RESULT 12
O36287 PRELIMINARY; PRT; 1255 AA.
AC O36287; O36301;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Structural polyprotein.
OS Venezuelan equine encephalitis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus; VEEV complex.
OX NCBI_TaxID=11036;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=66637, and 66457;
RX MEDLINE=97404681; PubMed=9261393;
RA Powers A.M., Oberste M.S., Brault A.C., Rico-Hesse R., Schmura S.M.,
RA Smith J.F., Kang W., Sweeney W.P., Weaver S.C.;
RT "Repeated emergence of epidemic/epizootic Venezuelan equine
encephalitis from a single genotype of enzootic subtype ID virus.";
RL J. Virol. 71:6697-6705(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=66637, and 66457;
RX MEDLINE=99214366; PubMed=10196323;
RA Wang E., Barrera R., Boshell J., Ferro C., Freier J.E., Navarro J.C.,
RA Salas R., Vasquez C., Weaver S.C.;
RT "Genetic and phenotypic changes accompanying the emergence of
epizootic subtype IC Venezuelan equine encephalitis viruses from an
enzootic subtype ID progenitor.";
RL J. Virol. 73:4266-4271(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=66637, and 66457;
RA Wang E., Weaver S.C.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF004458; AAC71187.2; -.

```

```

DR EMBL; AF004472; AAC71997.2; -.
DR HSSP; P05674; 1EPS.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002548; Alpha_E1_glycop.
DR InterPro; IPR000936; Alpha_E2_glycop.
DR InterPro; IPR002533; Alpha_E3_glycop.
DR InterPro; IPR000336; Flavi_glycoprote.
DR InterPro; IPR000930; Peptidase_S3.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00944; Alpha_core; 1.
DR Pfam; PF01589; Alpha_E1_glycop; 1.
DR Pfam; PF00943; Alpha_E2_glycop; 1.
DR Pfam; PF01563; Alpha_E3_glycop; 1.
DR PRINTS; PR00798; TOGAVIRIN.
KW Polyprotein.
SQ SEQUENCE 1255 AA; 138339 MW; 799350E33C807EAA CRC64;

Query Match 27.1%; Score 61; DB 2; Length 1255;
Best Local Similarity 47.6%; Pred. No. 97;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 11 DTVTESSDGLMNNQTLFLE 31
DB 758 ETIWESLDHLNNNQMFMIQ 778

RESULT 13
O7USQ0 PRELIMINARY; PRT; 3056 AA.
ID O7USQ0;
AC O7USQ0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Probable extracellular nuclease.
OS OrderedLocusNames=RB4375;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294140; CAD73746.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007154; P:cell communication; IEA.
DR InterPro; IPR006034; Asp/Glutamase.
DR InterPro; IPR003644; Calx_beta.
DR InterPro; IPR003368; Chlamydia_PMP.
DR InterPro; IPR002105; Dockerin_1.
DR InterPro; IPR006626; Pbh1.
DR InterPro; IPR011050; Pectin_lyase_like.
DR InterPro; IPR011506; Planc_extracel.
DR Pfam; PF03160; Calx-beta; 1.
DR Pfam; PF02415; Chlam_PMP; 59.
DR Pfam; PF04004; Dockerin_1; 2.
DR Pfam; PF07595; Planc_extracel; 1.
DR SMART; SM00710; Pbh1; 32.
DR PROSITE; PS00144; ASN_GLN_ASE_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 3056 AA; 296490 MW; A305EF2F431BF67 CRC64;

```

```
Query Match          26.9%; Score 60.5; DB 2; Length 3056;
Best Local Similarity 31.1%; Pred. No. 3.1e+02;
Matches 14; Conservative 6; Mismatches 14; Indels 11; Gaps 2;

QY 8 DGN-----DVTSSGLMNNNTQTFLFSLHSLTANTTKG 42
   ||| : ||||| : : : ||| :
Db 306 DGNADISGTTIRDTNFAAREGGGLMNNNTGT-MIVDGTILTGNVAG 349

RESULT 14
NRGL_CHICK
ID NRGL_CHICK STANDARD; PRT; 602 AA.
AC Q05199; O73750; O73751; O73752;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE (Pro-neuregulin-1 precursor (Pro-NRGL) [Contains: Neuregulin-1
DE (Acetylcholine receptor inducing activity) (ARIA)].
GN Name=NRGL; Synonyms=ARIA;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RC STRAIN=White leghorn; TISSUE=Brain;
RX MEDLINE=93201602; PubMed=8453670; DOI=10.1016/0092-8674(93)90407-H;
RA Falls D.L., Rosen K.M., Corfas G., Lane W.S., Fischbach G.D.;
RT "ARIA, a protein that stimulates acetylcholine receptor synthesis, is
RT a member of the neu ligand family.";
RL Cell 72:801-815 (1993).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3 AND 4).
RC TISSUE=Brain, and Spinal cord;
RX MEDLINE=98150951; PubMed=9491987; DOI=10.1016/S0896-6273(00)80454-7;
RA Yang X., Kuo Y., Devay P., Yu C., Role L.;
RT "A cysteine-rich isoform of neuregulin controls the level of
RT expression of neuronal nicotinic receptor channels during
RT synaptogenesis.";
RL Neuron 20:255-270 (1998).
CC -!- FUNCTION: Direct ligand for the ERBB tyrosine kinase receptors.
CC The multiple isoforms perform diverse functions: cysteine-rich
CC domain containing isoforms (isoforms 2-4) probably regulate the
CC expression of nicotinic acetylcholine receptors at developing
CC interneuronal synapses. The Ig-NRG isoform is required for the
CC initial induction and/or maintenance of the mature levels of
CC acetylcholine receptors at neuromuscular synapses.
CC -!- SUBCELLULAR LOCATION: Exists as a type I membrane protein and as a
CC proteolytically released soluble growth factor form. The membrane-
CC bound form does not seem to be active (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Comment=Additional isoforms seem to exist;
CC Name=1; Synonyms=ARIA, IG-NRG;
CC IsoId=Q05199-1; Sequence=Displayed;
CC Note=Contains an Ig-like domain;
CC Name=2; Synonyms=CRD-NRG-BETA1;
CC IsoId=Q05199-2; Sequence=VSP_003445;
CC Note=The EGF-like domain is replaced by a cysteine-rich domain
CC (CRD);
CC Name=3; Synonyms=CRD-NRG-BETA2A;
CC IsoId=Q05199-3; Sequence=VSP_003445, VSP_003446;
CC Note=The EGF-like domain is replaced by a cysteine-rich domain
CC (CRD);
CC Name=4; Synonyms=CRD-NRG-BETA2B;
CC IsoId=Q05199-4; Sequence=VSP_003445, VSP_003446, VSP_003447,
CC VSP_003448;
CC Note=The EGF-like domain is replaced by a cysteine-rich domain
CC (CRD);
CC -!- DEVELOPMENTAL STAGE: Isoforms 2-4 are detected at embryonic day 4
CC (ED4) in both visceral and somatic motor neurons of spinal cord
CC and is highest at ED6. Isoform 1 is not expressed until ED 6 in
```

Search completed: June 29, 2005, 09:01:43
Job time : 17.833 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:46:27 ; Search time 25.4145 Seconds
(without alignments)
635.503 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225

Sequence: 1 PLYSYTDGNTVTSSDGL.....NNQTQFLFLEHLLTANTTKG 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 487488

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgm2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgm2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgm2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgm2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgm2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgm2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgm2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgm2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 9: /cgm2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 10: /cgm2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 11: /cgm2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 12: /cgm2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 13: /cgm2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 14: /cgm2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 15: /cgm2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 16: /cgm2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 17: /cgm2_6/ptodata/1/pubpaa/US10F_PUBCOMB.pep.*
- 18: /cgm2_6/ptodata/1/pubpaa/US10G_PUBCOMB.pep.*
- 19: /cgm2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgm2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgm2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgm2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	44.0	18	17	US-10-718-321-1
2	48	21.3	44	15	US-10-424-599-190016
3	48	21.3	48	15	US-10-424-599-282496
4	47.5	21.1	40	15	US-10-424-599-264613
5	46	20.4	44	15	US-10-424-599-240616
6	45	20.0	49	15	US-10-424-599-179504
7	43.5	19.3	31	15	US-10-424-599-284397
8	43.5	19.3	42	16	US-10-437-963-166786
9	43	18.1	43	9	US-09-925-300-1031
10	42.5	18.9	43	16	US-10-425-115-191318
11	42	18.7	35	15	US-10-424-599-216152

12	42	18.7	42	15	US-10-424-599-265756	Sequence 265756,
13	42	18.7	42	15	US-10-424-599-277738	Sequence 277738,
14	42	18.7	50	16	US-10-425-115-367358	Sequence 367358,
15	41.5	18.4	37	15	US-10-424-599-269343	Sequence 269343,
16	41	18.2	32	16	US-10-425-115-328576	Sequence 328576,
17	41	18.2	37	16	US-10-425-115-330334	Sequence 330334,
18	41	18.2	39	16	US-10-425-115-194419	Sequence 194419,
19	41	18.2	39	16	US-10-608-029-29	Sequence 29, Appl
20	40.5	18.0	34	16	US-10-425-115-333901	Sequence 333901,
21	40.5	18.0	36	15	US-10-424-599-264025	Sequence 264025,
22	40	17.8	14	14	US-10-193-477-75	Sequence 75, Appl
23	40	17.8	33	15	US-10-432-412-16	Sequence 16, Appl
24	40	17.8	40	16	US-10-425-115-186134	Sequence 186134,
25	40	17.8	42	9	US-09-185-908-6	Sequence 6, Appl
26	40	17.8	42	10	US-09-482-682-25	Sequence 25, Appl
27	40	17.8	49	10	US-09-824-438-9	Sequence 9, Appl
28	39.5	17.6	25	16	US-10-601-656-56	Sequence 56, Appl
29	39.5	17.6	25	17	US-10-873-015-69	Sequence 69, Appl
30	39.5	17.6	42	17	US-10-472-928-730	Sequence 730, App
31	39	17.3	30	16	US-10-716-029-275	Sequence 275, App
32	39	17.3	30	16	US-10-425-115-362870	Sequence 362870,
33	39	17.3	43	15	US-10-424-599-147295	Sequence 147295,
34	39	17.3	46	15	US-10-424-599-203943	Sequence 203943,
35	39	17.3	46	16	US-10-425-115-200490	Sequence 200490,
36	39	17.3	47	16	US-10-425-115-232897	Sequence 232897,
37	39	17.3	47	16	US-10-425-115-362227	Sequence 362227,
38	39	17.3	48	15	US-10-424-599-179725	Sequence 179725,
39	39	17.3	49	15	US-10-424-599-147420	Sequence 147420,
40	39	17.3	49	16	US-10-425-115-222102	Sequence 222102,
41	38.5	17.1	48	15	US-10-305-133-18	Sequence 18, Appl
42	38.5	17.1	49	16	US-10-425-115-324897	Sequence 324897,
43	38	16.9	28	14	US-10-121-799-2	Sequence 2, Appl
44	38	16.9	28	14	US-10-404-626-2	Sequence 2, Appl
45	38	16.9	28	20	US-11-033-209-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-718-321-1
; Sequence 1, Application US/10718321
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailey, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-1

Query Match 44.0%; Score 99; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 SSDGLNNNTQFLFLEHS 33
|||||

Db 1 SSDGLWNNNQTLFLEHS 18

RESULT 2

US-10-424-599-190016
; Sequence 190016, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 190016
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142603C.1.pep
US-10-424-599-190016

Query Match 21.3%; Score 48; DB 15; Length 44;

Best Local Similarity 37.9%; Pred. No. 54; Mismatches 7; Indels 4; Gaps 1;

Matches 11; Conservative 7; Mismatches 7; Indels 4; Gaps 1;

QY 12 TWTSSDGLWNNN-----QTQLFLEHSLLT 36

Db 2 TPTQSSLNFYHSNTLSLSETSFPIKHSLLT 30

RESULT 3

US-10-424-599-282496
; Sequence 282496, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 282496
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_97115C.1.pep
US-10-424-599-282496

Query Match 21.3%; Score 48; DB 15; Length 48;

Best Local Similarity 39.3%; Pred. No. 59; Mismatches 8; Indels 6; Gaps 1;

Matches 11; Conservative 3; Mismatches 8; Indels 6; Gaps 1;

QY 20 LWNNNQTLFLEHSLTA-----NTTK 41

Db 7 MWKNRQTKSFLERHVLPGMSOHRNLTK 34

RESULT 4

US-10-424-599-264613
; Sequence 264613, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 264613

; LENGTH: 40

; TYPE: PRT

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT3847_80967C.1.pep

US-10-424-599-264613

Query Match 21.1%; Score 47.5; DB 15; Length 40;

Best Local Similarity 40.0%; Pred. No. 56; Mismatches 12; Conservative 3; Mismatches 10; Indels 5; Gaps 1;

Matches 12; Conservative 3; Mismatches 10; Indels 5; Gaps 1;

QY 11 DTVTSSDGLWNNNQTLFLEHSLLTANTT 40

Db 7 ETHTELDSPMQNNLFI-----LTATTS 31

RESULT 5

US-10-424-599-240616
; Sequence 240616, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 240616
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59303C.1.pep
US-10-424-599-240616

Query Match 20.4%; Score 46; DB 15; Length 44;

Best Local Similarity 36.8%; Pred. No. 1e+02; Mismatches 5; Indels 10; Gaps 3;

Matches 14; Conservative 5; Mismatches 5; Indels 10; Gaps 3;

QY 1 PLYSYTTDNDTVTSSDGLWNNNQTLFLEHSLLTAN 38

Db 11 PLFIY-----YKSSSES-WNSKNTLNF-SHDAHTFN 38

RESULT 6

US-10-424-599-179504
; Sequence 179504, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_65460C.1.pep
US-10-437-963-166786

Query Match          19.3%; Score 43.5; DB 16; Length 42;
Best Local Similarity 47.8%; Pred. No. 2.1e+02;
Matches 11; Conservative 6; Mismatches 3; Indels 3; Gaps 2;

Qy    20 LNNNQTLF--LEHSLLTANTTK 11
      :| | |:| |:| |:| |:| |:|
Db    15 IWTILNITKLYLKHTLL--NSTK 35

RESULT 9
US-09-925-300-1031
; Sequence 1031, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1031
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (7)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1031

Query Match          19.1%; Score 43; DB 9; Length 43;
Best Local Similarity 40.9%; Pred. No. 2.5e+02;
Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

Qy    21 WNNNQTLFLFLEHSLLTANTTKG 42
      |:| |:| |:| |:| |:|
Db    21 WDYRHEQLHLVHMLLIVEEVKG 42

RESULT 10
US-10-425-115-191318
; Sequence 191318, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 191318
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_106061C.1.pep
US-10-425-115-191318

Query Match          18.9%; Score 42.5; DB 16; Length 43;
Best Local Similarity 29.0%; Pred. No. 2.9e+02;
```



```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 269343
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_85237C.1.pep
US-10-424-599-269343
```

```
Query Match      18.4%; Score 41.5; DB 15; Length 37;
Best Local Similarity 37.5%; Pred. No. 3.3e+02;
Matches 12; Conservative 3; Mismatches 16; Indels 1; Gaps 1;
```

```
OY      1 PLYSYTTDGNQDTVTSSDGLMNNQTLPLEH 32
      ||| | | | | | | | | | | | | | |
Db      4 PLYRVTLVSSDD-CKHLDLISNKQTHYSLRH 34
```

```
Search completed: June 29, 2005, 10:14:19
Job time : 26.4145 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:03:05 ; Search time 28.7151 Seconds
(without alignments)
565.693 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81
Perfect score: 225
Sequence: 1 PLYSTTGNDVFTSSDGL.....NNQTQLFLEHLLTANTTKG 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 938430

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	44.9	18	6	Aao26686 Human KIM
2	99	44.0	18	6	Aao26687 Human KIM
3	99	44.0	18	6	Aao26673 Monoclonal
4	90	40.0	18	6	Aao26688 Human KIM
5	82	36.4	18	6	Aao26685 Human KIM
6	48.5	21.6	40	8	Ad807403 Staphyloc
7	46	20.4	49	4	Aam92081 Human dig
8	45.5	20.2	40	2	Aaw25651 Rat cadhe
9	45.5	20.2	40	2	Aaw13116 Rat cadhe
10	44.5	19.8	31	2	Aar85818 Peptide r
11	44	19.6	20	2	Aaw08045 HIV pepi
12	43	19.1	43	3	Ad564453 Human pro
13	42.5	18.9	20	8	Ad195222 OSPF-rela
14	42	18.7	40	2	Aar43565 Rat cadhe
15	42	18.7	40	2	Aaw25648 Rat cadhe
16	42	18.7	40	2	Aaw13113 Rat cadhe
17	42	18.7	49	2	Aar83059 Human IL-
18	41.5	18.4	32	2	Aaw18789 Z peptide
19	41	18.2	46	4	Aam4315 Rat EST e
20	41	18.2	46	7	Adf60572 Human con
21	41	18.2	46	1	Adf59574 Human pol
22	41	18.2	49	7	Adf59574 Human pol
23	41	18.2	50	5	Abp11026 Human ORF
24	40.5	18.0	50	4	Abg30194 Human hum
25	40	17.8	14	7	Ade03490 BGS-3 asp

26	40	17.8	14	8	Ado10179 Human int
27	40	17.8	33	5	Aae28518 Neisseria
28	40	17.8	41	3	Aab34080 Human sec
29	40	17.8	42	7	Adf48735 Adenovira
30	39.5	17.6	30	2	Aar84975 Peptide r
31	39.5	17.6	42	6	Abu00797 S. pneumo
32	39	17.3	40	2	Aaw25650 Rat cadhe
33	39	17.3	40	2	Aaw13115 Rat cadhe
34	39	17.3	40	3	Aar20698 Polymeric
35	39	17.3	41	2	Aar85528 Sugar con
36	39	17.3	41	3	Aab34079 Human sec
37	39	17.3	49	3	Aag50726 Arabidops
38	38.5	17.1	30	2	Aar85014 Peptide r
39	38.5	17.1	31	2	Aar85810 Peptide r
40	38.5	17.1	48	6	Ada22738 Therapeut
41	38	16.9	21	2	Aay34178 Human pre
42	38	16.9	28	1	Aap70718 Equine in
43	38	16.9	28	7	Ade36036 Anaplasma
44	38	16.9	28	7	Ade72984 E. equi m
45	38	16.9	35	2	Aaw62593 Peptide h

ALIGNMENTS

RESULT 1
AAO26686
ID AAO26686 standard; peptide; 18 AA.
XX
AC AAO26686;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 14.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO2002398920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
(BIOJ) BIOGEN INC.
(GEO) GEN HOSPITAL CORP.
PI Bailly V, Bonventre J;
XX
XX WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Disclosure; Fig 1A; 42pp; English.
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
XX -binding polypeptide that inhibits proteolytic release of a soluble
XX kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
XX The antibody, antibody derivative or antigen-binding polypeptide is
XX useful for treating or preventing renal disease or injury, e.g. renal
XX cancer. The antibody is also useful for inhibiting shedding of the KIM-1
XX polypeptide. This sequence represents a human KIM-1 mucin domain related
XX peptide of the invention

```
SQ Sequence 18 AA;
Query Match 44.9%; Score 101; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 TDGNDVTVESSDGLWNNN 24
DB 1 TDGNDVTVESSDGLWNNN 18

RESULT 2
AAO26687
ID AAO26687 standard; peptide; 18 AA.
AC AAO26687;
XX
XX 20-MAR-2003 (first entry)
XX Human KIM-1 mucin domain related peptide, SEQ ID No 15.
DE
DE Human KIM-1 mucin domain related peptide, SEQ ID No 15.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
XX Synthetic.
OS
XX WO200298920-A1.
PN
XX 12-DEC-2002.
PD
XX 31-MAY-2002; 2002WO-US017402.
PF
XX 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
XX (BIOJ ) BIOGEN INC.
PA (GEO ) GEN HOSPITAL CORP.
XX
XX Bailly V, Bonventre J;
PI
XX WPI; 2003-156845/15.
DR
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Claim 4; Page 25; 42pp; English.
PS
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a binding epitope of the invention
XX
XX Sequence 18 AA;
Query Match 44.0%; Score 99; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 16 SSDGLWNNNTQFLFLEHS 33
DB 1 SSDGLWNNNTQFLFLEHS 18

RESULT 4
AAO26688
ID AAO26688 standard; peptide; 18 AA.
XX
XX AAO26688;
AC
XX
XX 20-MAR-2003 (first entry)
XX
XX Human KIM-1 mucin domain related peptide, SEQ ID No 16.
DE
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
XX Synthetic.
OS
XX WO200298920-A1.
PN
XX
```

```
SQ Sequence 18 AA;
Query Match 44.9%; Score 101; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7 TDGNDVTVESSDGLWNNN 24
DB 1 TDGNDVTVESSDGLWNNN 18

RESULT 2
AAO26687
ID AAO26687 standard; peptide; 18 AA.
AC AAO26687;
XX
XX 20-MAR-2003 (first entry)
XX Human KIM-1 mucin domain related peptide, SEQ ID No 15.
DE
DE Human KIM-1 mucin domain related peptide, SEQ ID No 15.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
XX Synthetic.
OS
XX WO200298920-A1.
PN
XX 12-DEC-2002.
PD
XX 31-MAY-2002; 2002WO-US017402.
PF
XX 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
XX (BIOJ ) BIOGEN INC.
PA (GEO ) GEN HOSPITAL CORP.
XX
XX Bailly V, Bonventre J;
PI
XX WPI; 2003-156845/15.
DR
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Disclosure; Fig 1A; 42pp; English.
PS
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
CC peptide of the invention
XX
XX Sequence 18 AA;
Query Match 44.0%; Score 99; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 9.6e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 16 SSDGLWNNNTQFLFLEHS 33
DB 1 SSDGLWNNNTQFLFLEHS 18

RESULT 3
AAO26673
ID AAO26673 standard; peptide; 18 AA.
XX
```

```

PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
XX
PR 04-JUN-2001; 2001US-0295907P.
XX
PA (BIOJ ) BIOGEN INC.
PA (GEOH ) GEN HOSPITAL CORP.
XX
PI Bailly V, Bonventre J;
XX
DR WPI; 2003-156845/15.
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Disclosure; Fig 1A; 42pp; English.
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
CC peptide of the invention
XX
SQ Sequence 18 AA;
Query Match 40.0%; Score 90; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.00014;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
CC peptide of the invention
XX
SQ Sequence 18 AA;
Query Match 36.4%; Score 82; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0015;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 PLYSYTTDGNQDVTVE 15
Db |||||||
XX 4 PLYSYTTDGNQDVTVE 18
XX
XX
RESULT 6
ADS07403
ID ADS07403 standard; protein; 40 AA.
XX
AC ADS07403;
XX
DT 04-NOV-2004 (first entry)
XX
DE Staphylococcus epidermis polypeptide seqid 6698.
XX
XX
KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system.
XX
XX Staphylococcus epidermidis.
XX
XX US2004147734-A1.
XX
PD 29-JUL-2004.
XX
PF 01-DEC-2003; 2003US-00724972.
XX
PR 08-NOV-1997; 97US-0064964P.
PR 13-AUG-1998; 98US-00134001.
PR 29-NOV-1999; 99US-00450969.
XX
XX (DOUC/) DOUCETTE-STAMM L.
PA (BUSH/) BUSH D.
XX
XX Doucette-Stamm L, Bush D;
XX
XX WPI; 2004-580138/56.
DR N-PSDB; ADS03631.
XX
XX New isolated polypeptide and encoding nucleic acid derived from
PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT treating an S. epidermidis bacterial infection.
XX
XX Claim 17; SEQ ID NO 6698; 741pp; English.
XX
XX The invention describes an isolated nucleic acid comprising a nucleotide
CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC given in the specification. Also described are: a recombinant expression
CC vector; a cell comprising a recombinant expression vector of (1);

```


DR WPI; 1996-432980/43.

XX Determining the likelihood of maternal transmission of HIV-1 to foetus -

PT by measuring maternal reactivity with specific gp120 and gp41 derived

PT peptide(s), also used for diagnosing HIV in infants.

XX

PS Disclosure; Col 59-60; 63pp; English.

XX

XX AAW08016-W08073 represent HIV peptides that can be used in the method of

CC the invention. The method of the invention is for determining whether or

CC not a mother will transmit HIV-1 to a foetus. The method comprises

CC incubating a sample from the HIV-infected mother, with a collection of

CC HIV peptides. The HIV peptides includes at least one gp41 derived peptide

CC (see AAW07918-W07928), and at least one HIV gp120 derived peptide (see

CC AAW07909-W07917). The number of peptides that react with the sample is

CC determined, and this number is compared with a standard that shows

CC pattern reactivity for a patient of transmission status. A non-

CC transmissible HIV sample is indicated if the test sample reacts with twice

CC as many peptides as the standard. The method detects the presence of

CC neutralising antibodies that protect against mother to infant

CC transmission of HIV. These sequences can also be used in vaccines to

CC protect against transmission. Antibodies against these sequences can be

CC used for passive immunisation, and to generate anti-idiotypic antibodies

CC for use in vaccines or active immunotherapy. (Updated on 16-OCT-2003 to

CC standardise OS field)

XX

SQ Sequence 20 AA;

Query Match 19.6%; Score 44; DB 2; Length 20;

Best Local Similarity 39.1%; Pred. No. 1.4e+02;

Matches 9; Conservative 4; Mismatches 2; Indels 8; Gaps 1;

QY 1 PLYSYTTDGNVTSSDGLWNN 23

Db ||::||::||::||::WNN 16

2 PLFNSTWNGNNT-----WNN 16

RESULT 12

AAB56453

ID AAB56453 standard; protein; 43 AA.

XX

AC AAB56453;

XX

DT 13-MAR-2001 (first entry)

XX

DE Human prostate cancer antigen protein sequence SEQ ID NO:1031.

XX

KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KW neuroprotective; cytostatic; cardioprotective; immunomodulatory; muscular;

KW vulnary; gastrointestinal; nephrotropic; antiinfective; gynaecological;

KW antibacterial; gene therapy; neural; immune; reproductive; renal;

KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KW wound; infectious disease.

XX

OS Homo sapiens.

XX

PN WO200055174-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US005988.

XX

PR 12-MAR-1999; 99US-0124270P.

XX

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

XX

PI Rosen CA, Ruben SM;

XX

DR WPI; 2000-587513/55.

DR N-PSDB; AAF15656.

XX

PT Prostate cancer associated gene sequences, referred to as prostate cancer

PT antigens, useful for treatment, prevention, and diagnosis of disorders

PT such as prostate cancer.

XX

PS Claim 11; Page 1460; 2338pp; English.

XX

CC AAF15566 to AAF16505 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytostatic,

CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,

CC nephrotropic, antiinfective, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to

CC AAB57303 represent sequences used in the exemplification of the present

CC invention

XX

SQ Sequence 43 AA;

Query Match 19.1%; Score 43; DB 3; Length 43;

Best Local Similarity 40.9%; Pred. No. 4.7e+02;

Matches 9; Conservative 1; Mismatches 12; Indels 0; Gaps 0;

QY 21 WNNNTQLEHSLTANTTKG 42

Db ||::||::||::||::

21 WDYRHEQLHVMLLIVEVKG 42

RESULT 13

ADI95222

ID ADI95222 standard; peptide; 20 AA.

XX

AC ADI95222;

XX

DT 04-NOV-2004 (first entry)

XX

DE OSPF-related SIV envelope (Env) peptide - SEQ ID 130.

XX

KW immune response; overlapping synthetic peptide formulation; OSPF;

KW immunostimulant; virucide; antibacterial; antiparasitic; cytostatic;

KW vaccine; viral; bacterial; parasitic infection; prion disease;

KW neoplastic; toxin; SIV; envelope; Env.

XX

OS Simian immunodeficiency virus.

XX

PN WO2004002415-A2.

XX

PD 08-JAN-2004.

XX

PF 27-JUN-2003; 2003WO-US020322.

XX

PR 27-JUN-2002; 2002US-0392718P.

XX

PA (DAND) DANA FARBER CANCER INST INC.

XX

PI Ruprecht RM, Jiang S;

XX

DR WPI; 2004-082868/08.

XX

XX Modulating an immune response, useful for treating immune disorders, e.g.

PT viral, bacterial and parasitic infections, prion diseases, or neoplastic

PT diseases, administering to a subject an overlapping synthetic peptide

PT formulation.

XX

PS Example 1; SEQ ID NO 130; 175pp; English.

XX

CC The invention relates to a novel method for modulating an immune response

CC comprising administering to a subject an overlapping synthetic peptide

CC formulation (OSPF) which comprises a combination of single chain peptides

CC corresponding to the amino acid sequence of a protein of interest. The

```
CC method of the invention has immunostimulant, virucide, antibacterial,
CC antiparasitic and cytostatic applications and may be useful during
CC vaccine production and for treating immune disorders including viral,
CC bacterial and parasitic infections, prion diseases, neoplastic diseases,
CC as well as providing protection against toxins. The current sequence is
CC that of the OSPF-related SIV envelope (Env) peptide of the invention.
XX
XX Sequence 20 AA;
SQ
Query Match 18.9%; Score 42.5; DB 8; Length 20;
Best Local Similarity 52.6%; Pred. No. 2.1e+02;
Matches 10; Conservative 2; Mismatches 6; Indels 1; Gaps 1;
QY 13 VTSSDGLWNNQTQLFLE 31
Db 1 VTESFDA-WNNTVTQAE 18
RESULT 14
AAR43565
ID AAR43565 standard; protein; 40 AA.
XX
AC AAR43565;
XX
DT 25-MAR-2003 (revised)
DT 19-JUN-1994 (first entry)
XX
DE Rat cadherin-7.
XX
KW Antibodies; binding.
XX
OS Rattus rattus.
XX
FN WO9321302-A1.
XX
PD 28-OCT-1993.
XX
PF 19-APR-1993; 93WO-US003681.
XX
PR 17-APR-1992; 92US-00872643.
XX
PA (DOHE-) DOHENY EYE INST.
XX
PI Suzuki S;
XX
WPI; 1993-351714/44.
DR N-PSDB; AAQ51223.
XX
XX DNA encoding cadherins (CH)6,7,9 and 10 antibodies to CH 5,6,7,9 and 10 -
XX useful for modulating binding and regulatory activities of cadherins.
XX Claim 14; Page 41; 99pp; English.
XX
XX Rat cadherins-7 was isolated from a rat brain and retina cDNA library
XX using degenerate PCR primers designed from highly conserved segments in
XX the cytoplasmic domain of mouse N-, E, and P cadherins. See also AAR43564
XX -7. (Updated on 25-MAR-2003 to correct FN field.)
XX
XX Sequence 40 AA;
SQ
Query Match 18.7%; Score 42; DB 2; Length 40;
Best Local Similarity 39.1%; Pred. No. 5.8e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 2 LYSYTTDGNITVTSSDGLWNNN 24
Db 2 LQTYAFEGNGSVAESLSLSDNS 24
RESULT 15
AAW25648
ID AAW25648 standard; protein; 40 AA.
XX
```

```
AC AAW25648;
XX
XX 25-MAR-2003 (revised)
DT 04-NOV-1997 (first entry)
XX
DE Rat cadherin-7 cytoplasmic domain.
XX
KW Human; cadherin; rat; calcium-dependent cell adhesion protein;
KW superfamily; cytoskeleton; eatenin; cancer.
XX
OS Rattus rattus.
XX
FH Key Location/Qualifiers
FT Misc-difference 40
FT /note= "Encoded by CTC"
XX
FN US5646250-A.
XX
XX 08-JUL-1997.
XX
XX 01-NOV-1994; 94US-00332638.
XX
PR 17-APR-1992; 92US-00872643.
PR 19-APR-1993; 93US-00049460.
XX
PA (DOHE-) DOHENY EYE INST.
XX
PI Suzuki S;
XX
WPI; 1997-362997/33.
DR N-PSDB; AAT85423.
XX
XX Human and rat cadherin polypeptide(s) - mediate cell-cell adhesion.
XX
XX Example 1; Col 27-28; 56pp; English.
XX
XX This sequence represents the cytoplasmic domain of rat cadherin-7. The
XX invention specifically provides details of human cadherin-5, -8, -11, -12
XX and -13, and rat cadherin-8, -5, -11 and -13. Cadherins are calcium-
XX dependent cell adhesion proteins. They are glycosylated integral membrane
XX proteins that have an N-terminal extracellular domain that determines
XX binding specificity, a hydrophobic membrane spanning region and a C-
XX terminal cytoplasmic domain, which is highly conserved among members of
XX the superfamily. The C-terminal domain interacts with the cytoskeleton
XX through eatenins and other cytoskeleton-associated proteins. The novel
XX cadherin proteins may be used in the analysis of the role of cadherins in
XX various cancers. Sequence analysis of the cadherin proteins also allows
XX investigation of the structure and function of cadherin. The cadherin
XX proteins may be isolated by using anti-cadherin antibodies. These
XX antibodies may also be used to modulate the activity of cadherin and to
XX determine the tissue specific distribution of cadherin proteins. Each
XX subclass of cadherins has a unique tissue distribution pattern. (Updated
XX on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 40 AA;
SQ
Query Match 18.7%; Score 42; DB 2; Length 40;
Best Local Similarity 39.1%; Pred. No. 5.8e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;
QY 2 LYSYTTDGNITVTSSDGLWNNN 24
Db 2 LQTYAFEGNGSVAESLSLSDNS 24
Search completed: June 29, 2005, 09:41:05
Job time : 31.7151 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:27:45 ; Search time 5.611 Seconds
(without alignments)
720.211 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81
Perfect score: 225
Sequence: 1 LYSYTTDGNVTVTSSDGL.....NNQTQFLFLESLLTANTTKG 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 11837

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43.5	19.3	49	2 S49365	Ig mu chain transmembrane region - bowfin
2	42.5	18.9	46	2 A49281	pol protein - simi
3	39	17.3	30	2 D82251	hypothetical prote
4	38	16.9	35	2 S30112	aminoglycoside N6'
5	37	16.4	34	2 S68880	ribulose-bisphosph
6	37	16.4	47	2 AF0464	hypothetical prote
7	37	16.4	48	2 S29150	superoxide dismuta
8	36.5	16.2	46	2 F49281	pol protein - simi
9	36	16.0	50	2 B49410	t-complex polypept
10	35.5	15.8	29	2 A32860	biotin-binding pro
11	35.5	15.8	37	2 AG0540	hypothetical prote
12	35	15.6	50	2 A61149	endo-1.4-beta-xyla
13	34.5	15.3	46	2 D49281	pol protein - simi
14	34	15.1	31	2 S49191	hypothetical prote
15	34	15.1	40	2 PQ0202	endo-1.4-beta-xyla
16	34	15.1	41	2 F82525	hypothetical prote
17	33.5	14.9	40	2 H95183	hypothetical prote
18	33	14.7	38	2 H72344	pilE protein - Nei
19	32.5	14.4	31	2 H82682	hypothetical prote
20	32.5	14.4	46	2 A49281	pol protein - simi
21	32.5	14.4	48	2 PL0090	Ig heavy chain V r
22	32	14.2	24	2 A36912	hypothetical prote
23	32	14.2	26	2 B53113	Lys-gingipain form
24	32	14.2	28	2 G32351	34K class B flagel
25	32	14.2	29	2 I50214	protein-tyrosine-p
26	32	14.2	31	2 A57001	endo-1.4-beta-xyla
27	32	14.2	41	2 S56120	collagenase type I
28	32	14.2	42	2 I37543	MHC HLA-DR-beta-1
29	32	14.2	45	2 D45731	gene 39.2 protein

30	31	13.8	16	2 F49039	T-cell receptor be
31	31	13.8	19	2 I49422	L-lactate dehydrog
32	31	13.8	23	2 S47188	T-cell receptor J-
33	31	13.8	26	2 F45087	cysteine proteinas
34	31	13.8	31	2 B82151	hypothetical prote
35	31	13.8	33	2 C83732	hypothetical prote
36	31	13.8	47	2 B97377	hypothetical prote
37	31	13.8	49	2 S11814	opacity protein (c
38	31	13.8	50	2 A99941	hypothetical prote
39	31	13.8	50	2 E85789	hypothetical prote
40	31	13.8	50	2 S78674	hypothetical prote
41	30.5	13.6	17	2 C37396	pollen allergen re
42	30.5	13.6	29	2 B43937	endo-1.4-beta-xyla
43	30.5	13.6	42	2 C42799	photosystem I chal
44	30.5	13.6	49	2 S26136	outer membrane pro
45	30	13.3	20	2 S03505	T-cell receptor al

ALIGNMENTS

RESULT 1

S49365
IG mu chain transmembrane region - bowfin
C:Species: Ania calva (bowfin)
C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Mar-1998
C:Accession: S49365
R:Wilson, M.R.; Ross, D.A.; Miller, N.W.; Clem, L.W.; Middleton, D.L.; Warr, G.W.
submitted to the EMBL Data Library, October 1994
A:Description: Alternate pre-mRNA splicing pathways in the production of membrane IgM he
A:Reference number: S49365
A:Accession: S49365
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-49 <WIL>
A:Cross-references: EMBL:X82070; NID:g557218; PID:g557219

Query Match 19.3%; Score 43.5; DB 2; Length 49;
Best Local Similarity 34.5%; Pred. No. 83;
Matches 10; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 2 LYSYTTDGNVTVTSSDGLMNNNQLEL 30
DB 5 VFQYVDQNG-IEYNGDSLWNTVCTFIFL 32

RESULT 2

A49281
pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Bab34) (fragment)
C:Species: simian T-cell lymphotropic virus type 1, STLV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C:Accession: A49281
R:Saksena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Matha
; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Poiesz, B.J.
Virology 198, 297-310, 1994
A:Title: Seroprevalence, molecular, and phylogenetic analyses of simian T-cell leukemia
A:Reference number: A49281; MUID:94082462; PMID:8259665
A:Accession: A49281
A:Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-46 <SAK>
A:Cross-references: UNIPROT:Q88390; GB:L20351; NID:g431689; PIDN:AAA16596.1; PID:g431690
C:Superfamily: pol polyprotein

Query Match 18.9%; Score 42.5; DB 2; Length 46;
Best Local Similarity 39.4%; Pred. No. 1e+02;
Matches 13; Conservative 3; Mismatches 14; Indels 3; Gaps 2;

QY 2 LYSYTTDGNVTVTES--SDGLMNNNQTLFLEH 32
DB 13 LKRYTDRPDLPMDNALSIGLWTFNHLNV-LTH 44

```
RESULT 3
D82251
hypothetical protein VC1030 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: D82251
R:Heidelbergl, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
R:Chardeson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, E.
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: D82251
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-30 <HEI>
A:Cross-references: UNIPROT:Q9KT75; GB:AE004184; GB:AE003852; NID:96655483; PIDN:AAF9418
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1030
A:Map position: 1

Query Match 17.3%; Score 39; DB 2; Length 30;
Best Local Similarity 61.5%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 28 LFLHSLTLANTTT 40
DB 12 LFYEHLLRLNVT 24

RESULT 4
S30112
aminoglycoside N6'-acetyltransferase (EC 2.3.1.82) - Klebsiella pneumoniae plasmid pCFFO
C:Species: Klebsiella pneumoniae
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-May-2000
C:Accession: S30112
R:Mabilat, C.; Lourencao-Vital, J.; Goussard, S.; Courvalin, P.
Mol. Gen. Genet. 235, 113-121, 1992
A:Title: A new example of physical linkage between Tn1 and Tn21: the antibiotic multiple
A:Reference number: S30112; MUID:93062798; PMID:1331747
A:Accession: S30112
A:Molecule type: DNA
A:Residues: 1-35 <NAB>
A:Cross-references: EMBL:X64525
C:Genetics:
A:Gene: aacA4
A:Genome: plasmid
C:Keywords: acyltransferase; antibiotic resistance; coenzyme A

Query Match 16.9%; Score 38; DB 2; Length 35;
Best Local Similarity 58.3%; Pred. No. 2.7e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 YSYTTDGNDDT 14
DB 24 YSIVTNSDSDVT 35

RESULT 5
S68880
ribulose-bisphosphate carboxylase (EC 4.1.1.39) large chain - kidney bean chloroplast (f
C:Species: Chloroplast Phaseolus vulgaris (kidney bean)
C>Date: 04-Dec-1997 #sequence_revision 04-Dec-1997 #text_change 09-Jul-2004
C:Accession: S68880
R:Yoshida, T.; Minamikawa, T.
Eur. J. Biochem. 238, 317-324, 1996
A:Title: Successive amino-terminal proteolysis of the large subunit of ribulose 1,5-bisph
A:Reference number: S68880; MUID:96283823; PMID:8681940
A:Accession: S68880
A:Molecule type: protein
A:Residues: 1-34 <YOS>
A:Cross-references: UNIPROT:Q7M2E2
```

```
C:Genetics:
A:Genome: chloroplast
A:Function:
A:Description: catalyzes the reaction of carbon dioxide with D-ribulose 1,5-bisphosphate
C:Superfamily: ribulose-bisphosphate carboxylase large chain
C:Keywords: Calvin cycle; carbon dioxide fixation; carbon-carbon lyase; carboxy-lyase; c
Query Match 16.4%; Score 37; DB 2; Length 34;
Best Local Similarity 47.1%; Pred. No. 3.5e+02;
Matches 8; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 5 YTTDGNDDVTTESSDGLW 21
DB 11 YTPDYETKDAESSGTGW 27

RESULT 6
AF0464
hypothetical protein YPO3815 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF0464
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0464
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-47 <KUR>
A:Cross-references: UNIPROT:Q8ZAI7; GB:AL590842; PIDN:CAC93282.1; PID:gl5981729; GSPDB:G
C:Genetics:
A:Gene: YPO3815

Query Match 16.4%; Score 37; DB 2; Length 47;
Best Local Similarity 39.1%; Pred. No. 5.1e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

QY 17 SDGLWNNNQTLFLEHSLTLTANT 39
DB 8 NNSLGYNNNVQVY--HFLDLSNT 28

RESULT 7
S29150
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) III, cytosolic - rice (fragment)
C:Species: Oryza sativa (rice)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: S29150; E61531
R:Kanematsu, S.; Asada, K.
Plant Cell Physiol. 31, 99-112, 1990
A:Title: Characteristic amino acid sequences of chloroplast and cytosol isozymes of CuZn
A:Reference number: S29146
A:Accession: S29150
A:Molecule type: protein
A:Residues: 1-48 <KAN>
A:Cross-references: UNIPROT:Q7M237
R:Kanematsu, S.; Asada, K.
Free Radic. Res. Commun. 12, 383-390, 1991
A:Title: Chloroplast and cytosol isozymes of CuZn-superoxide dismutase: their characteri
A:Accession: E61531
A:Molecule type: protein
A:Residues: 1-48 <KA2>
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase [Cu-Zn]
C:Keywords: copper; cytosol; metalloprotein; oxidoreductase; zinc
F:44,46/Binding site: copper (His) #status predicted

Query Match 16.4%; Score 37; DB 2; Length 48;
```

```

RESULT 10
A32860
biotin-binding protein I - chicken (fragment)
C;Species: Gallus gallus (chicken)
C;Date: 22-Nov-1989 #sequence_revision 22-Nov-1989 #text_change 24-Jun-1993
C;Accession: A32860
R;Bush, L.; White III, H.B.
J. Biol. Chem. 264, 5741-5745, 1989
A;Title: Conversion of domains into subunits in the processing of egg yolk biotin-binding
A;Reference number: A32860; MUID:89174628; PMID:2925632
A;Accession: A32860
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-29 <BUS>

Query Match 15.8%; Score 35.5; DB 2; Length 29;
Best Local Similarity 52.9%; Pred. No. 4.5e+02;
Matches 9; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 19 GLWNNNQTLFLEHSL 35
|||||4|||
DB 9 GLWENEQDSL-MEISAL 24

RESULT 11
AG0540
hypothetical protein STY0342 [imported] - Salmonella enterica subsp. enterica serovar Typhi
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0540
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0540
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-37 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08767.1; PID:G16501585; GSPDB:GN00176
C;Genetics:
A;Gene: STY0342

Query Match 15.8%; Score 35.5; DB 2; Length 37;
Best Local Similarity 61.5%; Pred. No. 6e+02;
Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 7 TDGNDTVTSSDG 19
|||||
DB 4 TDGNDN-TRKQDG 15

RESULT 12
A61149
endo-1,4-beta-xylanase (EC 3.2.1.8) - fungus (Trichoderma viride) (fragment)
C;Species: Trichoderma viride
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 09-Jul-2004
C;Accession: A61149
R;Ujile, M.; Roy, C.; Yaguchi, M.
Appl. Environ. Microbiol. 57, 1860-1862, 1991
A;Title: Low-molecular-weight xylanase from Trichoderma viride.
A;Reference number: A61149; MUID:91336731; PMID:1872616
A;Accession: A61149
A;Molecule type: protein
A;Residues: 1-50 <UJI>
A;Cross-references: UNIPROT:Q7M4V8
C;Function:
A;Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans
A;Pathway: xylan degradation
C;Superfamily: endo-1,4-beta-xylanase; endo-1,4-beta-xylanase homology

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

C;Keywords: glycosidase; hydrolase; polysaccharide degradation; pyroglutamic acid
F.1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 15.6%; Score 35; DB 2; Length 50;

Best Local Similarity 58.3%; Pred. No. 9.8e+02;

Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 YSYTTDGNDDTVT 14

|||||::|

Db 15 YSYWNDHGGVT 26

|||||::|

RESULT 13

D49281

pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Pat24) (fragment)

C;Species: simian T-cell lymphotropic virus type 1, STLV-1

C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004

C;Accession: D49281

R;Sakena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Math

; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Poiesz, B.J.

Virology 198, 297-310, 1994

A;Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukem

A;Reference number: A49281; PMID:94082462; PMID:8259665

A;Accession: D49281

A;Status: preliminary; translation not shown

A;Molecule type: DNA

A;Residues: 1-46 <SAK>

A;Cross-references: UNIPROT:Q88403; GB:L20353; NID:9431711; PIDN:AAA16557.1; PID:9431712

C;Superfamily: pol polyprotein

Query Match

Best Local Similarity 15.3%; Score 34.5; DB 2; Length 46;

Matches 12; Conservative 3; Mismatches 15; Indels 3; Gaps 2;

QY 2 LYSYTTDGNDDTVTES--SDGLMNNNQTLFLEH 32

|||||::|

Db 13 LYKYFTDKPDLPMDNALSITLTINHLNV-LTH 44

|||||::|

RESULT 14

S49191

hypothetical protein 6 - Azotobacter vinelandii (fragment)

C;Species: Azotobacter vinelandii

C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 09-Jul-2004

C;Accession: S49191

R;Wientjens, R.; van Dongen, W.; Haaker, H.

submitted to the EMBL Data Library, April 1992

A;Description: Molecular cloning of fixA, fixB, fixC and fixX genes of Azotobacter vinel

A;Reference number: S49186

A;Accession: S49191

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-31 <WIE>

A;Cross-references: UNIPROT:Q44509; EMBL:X65515

Query Match

Best Local Similarity 15.1%; Score 34; DB 2; Length 31;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 24 NQTQLFLEHSL 34

|||||::|

Db 2 NDIQLFLAHAI 12

|||||::|

RESULT 15

P00202

endo-1,4-beta-xylanase (EC 3.2.1.8) II - Nocardiopsis dassonvillei (fragment)

N;Alternate names: xylanase

C;Species: Nocardiopsis dassonvillei

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004

C;Accession: P00202

R;Tsujibo, H.; Sakamoto, T.; Miyamoto, K.; Hasegawa, T.; Fujimoto, M.; Inamori, Y.

Agric. Biol. Chem. 55, 2173-2174, 1991

A;Title: Amino acid compositions and partial sequences of xylanases from a new subspecies
A;Reference number: P00202; PMID:92118324; PMID:1368735

A;Accession: P00202

A;Molecule type: protein

A;Residues: 1-40 <TSU>

A;Cross-references: UNIPROT:Q9R5P8

A;Experimental source: subsp. alba OPC-18

C;Comment: This enzyme is important for the breakdown of xylan.

C;Function:

A;Description: catalyzes the hydrolysis of 1,4-beta-xylosidic linkages in xylans

A;Pathway: xylan degradation

A;Superfamily: endo-1,4-beta-xylanase; endo-1,4-beta-xylanase homology

C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 15.1%; Score 34; DB 2; Length 40;

Best Local Similarity 50.0%; Pred. No. 1e+03;

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 YSYTTDGNDDTVT 14

|||||::|

Db 16 YSFWDGGGSVS 27

|||||::|

Search completed: June 29, 2005, 09:47:29

Job time : 6.611 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:03:05 ; Search time 55.3792 Seconds
(without alignments)
565.693 Million cell updates/sec

Title: US-10-718-321-6
Perfect score: 432
Sequence: 1 MFLPRQHPEVATSPSSPOP.....NNQTQLFLEHSLTANTTKG 81

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 938430

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	23.4	18	6	AAO26686 Human KIM
2	99	22.9	18	6	AAO26687 Human KIM
3	99	22.9	18	6	AAO26673 Monoclonal
4	98	22.7	18	6	AAO26681 Human KIM
5	97	22.5	18	6	AAO26682 Human KIM
6	96	22.2	18	6	AAO26683 Human KIM
7	95	22.0	18	6	AAO26685 Human KIM
8	92	21.3	18	6	AAO26684 Human KIM
9	90	20.8	18	6	AAO26688 Human KIM
10	60	13.9	50	4	AAO20331 Peptide #
11	60	13.9	50	4	ABO40865 Peptide #
12	60	13.9	50	4	AAO34632 Peptide #
13	60	13.9	50	4	ABO25019 Protein #
14	60	13.9	50	4	AAO74520 Human bon
15	60	13.9	50	4	AAO61724 Human bra
16	60	13.9	50	4	ABO56310 Human liv
17	60	13.9	50	5	ABO44381 Human pep
18	54.5	12.6	40	6	ABO91655 P. papata
19	53	12.3	39	4	ABO53259 Human tes
20	52.5	12.2	47	3	ABO56300 Human sec
21	52	12.0	47	8	ABO59205 Human gen
22	51.5	11.9	50	4	ABO50592 Human sec
23	51.5	11.9	50	6	ABO44849 Novel hum
24	51.5	11.9	50	7	ABO26329 Protein a
25	50	11.6	35	4	AAO2139 Mouse alp

26	50	11.6	35	5	AAU99631	Aau99631 Mouse sec
27	49	11.3	50	2	AAU12405	Aay12405 Human 5'
28	48.5	11.2	40	2	AAR88860	Aar88860 Peptide t
29	48.5	11.2	40	8	ADS07403	Ads07403 Staphyloc
30	48	11.1	19	2	AAR75547	Aar75547 P. Commun
31	48	11.1	20	2	AAO80845	Aao80845 HIV pepti
32	48	11.1	48	8	ABO55454	Abos5454 Human gen
33	48	11.1	50	4	AAU44389	Aau44389 Propionib
34	48	11.1	50	6	ABM40908	Abm40908 Human bre
35	47.5	11.0	46	4	AAO63400	Aao63400 Human bre
36	47.5	11.0	50	5	ABP35345	Abp35345 Human ORF
37	47	10.9	23	6	ABR91640	Abri91640 P. papata
38	46.5	10.8	38	2	AAU11409	Aay11409 Human 5'
39	46.5	10.8	46	4	AAO96710	Aao96710 Human rep
40	46.5	10.8	46	4	AAU22911	Aau22911 Human pro
41	46.5	10.8	46	7	ADJ09484	Adj09484 Human pro
42	46	10.6	20	2	AAW45839	Aaw45839 Peptide r
43	46	10.6	26	2	AAV31378	Aay31378 Human pro
44	46	10.6	30	6	AAE34173	Aae34173 T-cell st
45	46	10.6	38	8	ADP73905	Adp73905 Human ant

ALIGNMENTS

RESULT 1
AAO26686
ID AAO26686 standard; peptide; 18 AA.
XX
AC AAO26686;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 14.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
(BIOJ) BIOGEN INC.
(GEO) GEN HOSPITAL CORP.
PI Bailly V, Bonventre J;
XX
XX WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Disclosure; Fig 1A; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
XX peptide of the invention

```
SQ Sequence 18 AA;
  Query Match      23.4%; Score 101; DB 6; Length 18;
  Best Local Similarity 100.0%; Pred. No. 0.002;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 TDGNDVTVTSSDGLWNNN 63
  |||||
Db 1 TDGNDVTVTSSDGLWNNN 18

RESULT 2
AAO26687
ID AAO26687 standard; peptide; 18 AA.
XX
AC AAO26687;
XX
DT 20-MAR-2003 (first entry)
DE Human KIM-1 mucin domain related peptide, SEQ ID No 15.
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 15.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
XX
PR 04-JUN-2001; 2001US-0295907P.
XX
PA (BIOJ ) BIOGEN INC.
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Bailly V, Bonventre J;
XX
XX WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Claim 4; Page 25; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a binding epitope of the invention
XX
SQ Sequence 18 AA;
  Query Match      22.9%; Score 99; DB 6; Length 18;
  Best Local Similarity 100.0%; Pred. No. 0.0032;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 SSDGLWNNNTQFLFLEHS 72
  |||||
Db 1 SSDGLWNNNTQFLFLEHS 18

RESULT 4
AAO26681
ID AAO26681 standard; peptide; 18 AA.
XX
AC AAO26681;
XX
DT 20-MAR-2003 (first entry)
DE Human KIM-1 mucin domain related peptide, SEQ ID No 9.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
```



```

CC peptide of the invention
XX Sequence 18 AA;
SQ
  Query Match      22.2%; Score 96; DB 6; Length 18;
  Best Local Similarity 100.0%; Pred. No. 0.0065;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QPAETHPTTLOGAIRREP 36
Db 1 QPAETHPTTLOGAIRREP 18
  |||||
  |||||

RESULT 7
AAO26685
ID AAO26685 standard; peptide; 18 AA.
XX
AC AAO26685;
XX
XX 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 13.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
XX WO200298920-A1.
FN
XX
PD 12-DEC-2002.
XX
XX 31-MAY-2002; 2002WO-US017402.
XX
XX 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
XX (BIOJ ) BIOGEN INC.
PA (GEO ) GEN HOSPITAL CORP.
XX
XX Bailly V, Bonventre J;
PI
XX
XX WPI; 2003-156845/15.
DR
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Disclosure; Fig 1A; 42pp; English.
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
CC peptide of the invention
XX
XX Sequence 18 AA;
PS
SQ
  Query Match      21.3%; Score 92; DB 6; Length 18;
  Best Local Similarity 100.0%; Pred. No. 0.017;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 28 LQGAIRREPTSSPLYSYT 45
Db 1 LQGAIRREPTSSPLYSYT 18
  |||||
  |||||

RESULT 9
AAO26688
ID AAO26688 standard; peptide; 18 AA.
XX
AC AAO26688;
XX
XX 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 16.
XX
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.

```

XX PN WO200298920-A1.
 XX PD 12-DEC-2002.
 XX PF 31-MAY-2002; 2002WO-US017402.
 XX PR 01-JUN-2001; 2001US-0295449P.
 XX PR 04-JUN-2001; 2001US-0295907P.
 XX PA (BIOJ) BIOGEN INC.
 XX PA (GHEO) GEN HOSPITAL CORP.
 XX PI Bailly V, Bonventre J;
 XX WI WIPI; 2003-156845/15.
 XX New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX Disclosure; Fig 1A; 42pp; English.
 XX The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX Sequence 18 AA;
 SQ Query Match 20.8%; Score 90; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.027;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 64 QTQLFLEHSLTANTTKG 81
 Db 1 QTQLFLEHSLTANTTKG 18
 RESULT 10
 AAM20331
 ID AAM20331 standard; protein; 50 AA.
 AC AAM20331;
 XX 12-OCT-2001 (first entry)
 XX Peptide #6765 encoded by probe for measuring cervical gene expression.
 DE Probe; human; microarray; gene expression; cervical epithelial cell;
 XX cervical cancer.
 KW Homo sapiens.
 XX WO200157278-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000670.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WIPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.
 XX Claim 27; SEQ ID NO 33500; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WIPI; 2001-488901/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX Claim 27; SEQ ID NO 25157; 487pp; English.
 XX The present invention relates to human single exon nucleic acid probes
 CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human Hela cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 50 AA;
 SQ Query Match 13.9%; Score 60; DB 4; Length 50;
 Best Local Similarity 44.4%; Pred. No. 1.1e+02;
 Matches 16; Conservative 3; Mismatches 9; Indels 8; Gaps 2;
 QY 12 ATSPSS-----PQPAETHPTTLOGAIRREPTSSP 40
 Db 5 ATSPSSRAAAPTUPAPPTVAPTACPGAVPR-PTTTP 39
 RESULT 11
 ABB40865
 ID ABB40865 standard; peptide; 50 AA.
 XX AC ABB40865;
 XX 04-FEB-2002 (first entry)
 XX Peptide #8371 encoded by human foetal liver single exon probe.
 DE Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX Homo sapiens.
 XX WO200157277-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000669.
 XX 04-FEB-2000; 2000US-0180312P.
 XX 26-MAY-2000; 2000US-0207456P.
 XX 30-JUN-2000; 2000US-00608408.
 XX 03-AUG-2000; 2000US-00632366.
 XX 21-SEP-2000; 2000US-0234687P.
 XX 27-SEP-2000; 2000US-0236359P.
 XX 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WIPI; 2001-483447/52.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human fetal liver.
 XX Claim 27; SEQ ID NO 33500; 639pp + Sequence Listing; English.
 XX The invention relates to a single exon nucleic acid probe for measuring


```

AC AM74520;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 34826.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-489900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX Example 4; SEQ ID NO 34826; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention
XX
XX Sequence 50 AA;
XX
Query Match 13.9%; Score 60; DB 4; Length 50;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 16; Conservative 3; Mismatches 9; Indels 8; Gaps 2;

QY 12 ATSPSS-----PQPAETHPTTLOGAIRREPTSSP 40
Db ||||| | | | | | | | | | | | | | | | |
5 ATSPSSRAAPTLPAPPTVAPTACPGAVPR-PTTTP 39

RESULT 15
ANM61724.
ID ANM61724 standard; protein; 50 AA.
XX
AC ANM61724;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 33829.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX

```

```

PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
XX Example 4; SEQ ID NO 33829; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention
XX
XX Sequence 50 AA;
XX
Query Match 13.9%; Score 60; DB 4; Length 50;
Best Local Similarity 44.4%; Pred. No. 1.1e+02;
Matches 16; Conservative 3; Mismatches 9; Indels 8; Gaps 2;

QY 12 ATSPSS-----PQPAETHPTTLOGAIRREPTSSP 40
Db ||||| | | | | | | | | | | | | | | | |
5 ATSPSSRAAPTLPAPPTVAPTACPGAVPR-PTTTP 39

Search completed: June 29, 2005, 09:41:02
Job time : 58.3792 Secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:27:45 ; Search time 10.8212 Seconds
(without alignments)
720.211 Million cell updates/sec

Title: US-10-718-321-6
Perfect score: 432
Sequence: 1 MPLPRQNHPEVATSPSPQP.....NNQTQLFLESLLTANTTKG 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 11837

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	13.2	42	2 J70082	glycoprotein Ib al
2	51	11.8	43	2 JU0404	hypothetical 5.2K
3	47.5	11.0	46	2 S07073	arabinogalactan pr
4	47.5	11.0	47	2 AE3429	hypothetical prote
5	43.5	10.1	49	2 S49385	ig mu chain transm
6	43	10.0	31	2 T29628	hypothetical prote
7	43	10.0	35	2 S30112	aminoglycoside N6'
8	43	10.0	42	2 T07030	extensin - tomato
9	42.5	9.8	46	2 A49281	pol protein - simi
10	42	9.7	41	2 C87603	hypothetical prote
11	40.5	9.4	25	2 I40692	cenA protein (IgA)
12	40.5	9.4	32	2 B38839	histone H1 - sea u
13	40.5	9.4	50	2 AH2135	hypothetical prote
14	39.5	9.1	47	2 A40703	androgen-binding p
15	39	9.0	30	2 D82251	hypothetical prote
16	39	9.0	42	2 A99578	lipoprotein vsaH [
17	38	8.8	36	2 A37172	collagen alpha 1(X
18	38	8.8	48	2 B84436	hypothetical prote
19	37	8.6	31	2 B82151	hypothetical prote
20	37	8.6	34	2 S68880	ribulose-bisphosph
21	37	8.6	44	2 A19434	hypothetical prote
22	37	8.6	45	2 S24713	ig alpha chain, tr
23	37	8.6	47	2 AF0464	hypothetical prote
24	37	8.6	48	2 S29150	superoxide dismuta
25	37	8.6	50	2 A29789	mucin - sheep (fra
26	36.5	8.4	22	2 S68616	histone H1 - sea u
27	36.5	8.4	36	2 S12384	pol protein - simi
28	36.5	8.4	46	2 F49281	sepiapterin reduct
29	36.5	8.4	47	2 S32108	

troponin T 2fa - r
androgen-binding p
phospholipase A2 (
hypothetical prote
coagulation factor
t-complex polypept
outer membrane pro
biotin-binding pro
hypothetical prote
histone H5 - pigeo
hypothetical prote
gamma-hordein 3 -
beta-fructofuranos
ATP synthase beta-
translation initia
hypothetical prote

ALIGNMENTS

RESULT 1

I70082
glycoprotein Ib alpha variant B - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C;Accession: I70082
R;Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.
J. Biol. Chem. 267, 10055-10061, 1992
A;Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of t
ations.
A;Reference number: I55355; MUID:92250564; PMID:1577776
A;Accession: I70082
A;Status: preliminary; translated from GE/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-42 <RES>
A;Cross-references: UNIPROT:Q16469; GB:S34439; NID:G249178; PIDN:AAB22153.1; PID:G249179
C;Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein r
C;Keywords: glycoprotein

Query Match 13.2%; Score 57; DB 2; Length 42;
Best Local Similarity 43.9%; Pred. No. 26;
Matches 18; Conservative 3; Mismatches 10; Indels 10; Gaps 3;

Qy 9 EPVATSPSPQPAE---THPTTLQGAIRREPTSPLYSYTT 46
Db 5 EP-APSPTEPTSEPAPSPPT-----PEPTSEPAPSPPT 38

RESULT 2

JU0404
hypothetical 5.2K protein - clover yellow mosaic virus
C;Species: clover yellow mosaic virus
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C;Accession: JU0404
R;Abouhaider, M.G.; Lai, R.
J. Gen. Virol. 70, 1871-1875, 1989
A;Title: Nucleotide sequence of the 3'-terminal region of clover yellow mosaic virus RNA
A;Reference number: JU0402; MUID:89293092; PMID:2738582
A;Accession: JU0404
A;Molecule type: genomic RNA
A;Residues: 1-43 <ABO>
A;Cross-references: UNIPROT:P16485; GB:D00485; NID:G221218; PIDN:BAA00374.1; PID:d100082
A;Note: this protein is encoded by sequence in front of the coat protein-encoding sequen

Query Match 11.8%; Score 51; DB 2; Length 43;
Best Local Similarity 45.8%; Pred. No. 1e+02;
Matches 11; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

Qy 4 PRONHEPVATSPSPQPAETHPTT 27
Db 20 PRQSHLPAQT-----PQRLSYPKT 39

R;van der Bycken, W.; de Almeida Engler, J.; Inze, D.; van Montagu, M.; Gheysen, G.
Plant J. 9, 45-54, 1996
A;Title: A molecular study of Meloidogyne incognita-induced feeding sites.
A;Reference number: Z15873; MUID:96158345; PMID:8580972
A;Accession: T07030
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-42 <VAN>
A;Cross-references: UNIPROT:Q40145; EMBL:Z46675; NID:G575949; PIDN:CAA86660.1; PID:G575949
A;Experimental source: strain Marmande; Meloidogyne incognita-induced root knot
C;Superfamily: hydroxyproline-rich glycoprotein
C;Keywords: cell wall; glycoprotein; hydroxyproline

Query Match 10.0%; Score 43; DB 2; Length 42;
Best Local Similarity 29.7%; Pred. No. 6.1e+02;
Matches 11; Conservative 2; Mismatches 12; Indels 12; Gaps 1;

Qy 10 PVATSPSSPQPAETHPTTLQGAIRREPTSSPLXYTT 46
Db 11 PIYKSPPTPAYNSP-----PPYDYLYTS 35

RESULT 9
A49281
pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Bab34) (fragment)
C;Species: simian T-cell lymphotropic virus type 1, STLV-1
C;Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C;Accession: A49281
R;Saksena, N.K.; Herve, V.; Durand, J.P.; Lequanno, B.; Diop, O.M.; Digoutte, J.P.; Math
; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Potesz, B.J.
Virology 198, 297-310, 1994
A;Title: Seroprevalence, molecular, and phylogenetic analyses of simian T-cell leuko
A;Reference number: A49281; MUID:94082462; PMID:8259665
A;Accession: A49281
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 1-46 <SAK>
A;Cross-references: UNIPROT:Q88390; GB:L20351; NID:G431689; PIDN:AAA16596.1; PID:G431690
C;Superfamily: pol polyprotein

Query Match 9.8%; Score 42.5; DB 2; Length 46;
Best Local Similarity 39.4%; Pred. No. 7.6e+02;
Matches 13; Conservative 3; Mismatches 14; Indels 3; Gaps 2;

Qy 41 LYSYTTDGNVTVES--SDGLWNNQTLFLEH 71
Db 13 LYKYFTDKPDLPMDNALSIGLWTFNLNV-LTH 44

RESULT 10
C87603
hypothetical protein CC2859 [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: C87603
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolof
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: C87603
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-41 <STO>
A;Cross-references: UNIPROT:Q9A4H5; GB:AE005673; NID:gl3424473; PIDN:AAK24823.1; GSPDB:G
A;Gene: CC2859

Query Match 9.7%; Score 42; DB 2; Length 41;
Best Local Similarity 29.6%; Pred. No. 7.4e+02;
Matches 8; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

Qy 4 PRNHPEVATSPSSPQPAETHPTTLQ 30
Db 12 PKGLNDRSTGGQPPRPRIIDPDALRG 38

RESULT 11
I40692
cena protein (Iga1h) - Cellulomonas fimi (fragment)
C;Species: Cellulomonas fimi
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 22-Oct-1999
C;Accession: I40692
R;Miller, P.B.; Shen, H.; Gilkes, N.R.; Kilburn, D.G.; Miller, R.C.; Plaut, A.G.; Warren
FEMS Microbiol. Lett. 92, 199-204, 1992
A;Title: Endoglucanase A from Cellulomonas fimi in which the hinge sequence of human Iga
hoese.
A;Reference number: I40692
A;Accession: I40692
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-25 <RES>
A;Cross-references: EMBL:X65780; NID:g312035; PIDN:CAA46663.1; PID:g312036

Query Match 9.4%; Score 40.5; DB 2; Length 25;
Best Local Similarity 32.4%; Pred. No. 5.7e+02;
Matches 12; Conservative 1; Mismatches 7; Indels 17; Gaps 1;

Qy 10 PVATSPSSPQPAETHPTTLQGAIRREPTSSPLXYTT 46
Db 6 PPTSPSTP-----PTPSPLQSNVT 25

RESULT 12
B38839
histone H1 - sea urchin (Strongylocentrotus intermedius) (fragment)
C;Species: Strongylocentrotus intermedius
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: B38839
R;Suzuki, M.; Sugitara, M.; Ebashi, S.
J. Biochem. 108, 347-355, 1990
A;Title: Sea urchin protease specific to the SPKK motif in histone.
A;Reference number: PS0145; MUID:91115778; PMID:2126010
A;Accession: B38839
A;Molecule type: protein
A;Residues: 1-32 <SUZ>
A;Cross-references: UNIPROT:Q7M407
C;Superfamily: histone H1
C;Keywords: chromosomal protein; nucleosome

Query Match 9.4%; Score 40.5; DB 2; Length 32;
Best Local Similarity 38.5%; Pred. No. 7.7e+02;
Matches 10; Conservative 1; Mismatches 8; Indels 7; Gaps 1;

Qy 15 PSSPQPAETHPTTLQGAIRREPTSSP 40
Db 1 PASPQKRAASP-----RSPKKSP 19

RESULT 13
AH2135
hypothetical protein asl2639 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C;Accession: AH2135
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AH2135
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <KUR>

A; Cross-references: UNIPROT:Q8VTS4; GB:BA000019; PIDN:BA874338.1; PID:gl17131732; GSPDB:G
A; Experimental source: strain PCC 7120
C; Genetics:
A; Gene: *agl2639*

Query Match 9.4%; Score 40.5; DB 2; Length 50;
Best Local Similarity 27.0%; Pred. No. 1.3e+03;
Matches 10; Conservative 5; Mismatches 19; Indels

Qy 7 NHEPVATSPSSPQPAETHP---TTLQGAIRREPTSSP 40
 10 NNTKTNNPPTPKPPKLLHPKLOPTHPNLQKNSPTKSP 46

Db

RESULT 14

A40703
androgen-binding protein ABP (N-terminal, exon A) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A40703
R:Sullivan, P.M.; Wang, Y.M.; Joseph, D.R.
Mol. Endocrinol. 7, 702-715, 1993
A:Title: Identification of an alternate promoter in the rat androgen-binding protein/sex
tion.

A:Reference number: A40703; **MUID:**93302779; **PMID:**7686253

A;Accession: A40703
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-47 <SUL>
A;Cross-references: UNIPROT:Q
A;Note: sequence extracted from

Query Match 9.1%; Score 39.5; DB 2; Length 47;
Best Local Similarity 42.3%; Pred. No. 1.5e+03;
Matches 11; Conservative 3; Mismatches 11; Indels

Qy 11 V A T S P S S - P Q P A E T H P T T L Q G A I R R E 35
 | | : | : | : | : | : | : | : | :
Dd 6 V P L Q P S T V P V F S F P R P T T L E A N Q S R E 31

RESULT 15

hypothetical protein VC1030 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
C/Species: *Vibrio cholerae*
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: D8251
R/Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
L, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.N.
Nature 406, 477-483, 2000
A/Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A/Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82251

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-30 <HEI>

A;Cross-references: UNIPROT:Q9KT75; GB:AE004184; GB:AE003852; NID:

A; Experimental source: serogroup O1; strain N16961; prototype El Tor C. Genetica.

C;Gene: VC1030

A;Map position: 1

100

Query Match	9.0%	Score 39;	DB 2;	Length 30;
-------------	------	-----------	-------	------------

Best Local Similarity 61.5%; Pred. No. 9.9e+02;

Matches 8; Conservative 0; Mismatches 5

67 I.FI.FHSI.I.TANNT 79

Qy 67 LF 6HSLTIANI 173

Db 12 LFYEHELLRLNVT 24

Search completed: June 29, 2005, 09:47:28


```

[1]
RP SEQUENCE FROM N.A.
RA Lee C., Lee S.-H., Han E., Kim S.-J.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY280886; AAP43056.1; -.
DR HSSP; P03277; 1P2Z.
DR GO; GO:0019028; C:viral capsid; IEA.
DR InterPro; IPR000736; F:structural molecule activity; IEA.
DR Pfam; PF01065; Adeno_hexon; 1.
DR ProDom; PD002815; Adeno_hexon; 1.
FT NON_TER 1
FT NON_TER 30
SQ SEQUENCE 30 AA; 3450 MW; 531F6F951E27039B CRC64;

Query Match 11.1%; Score 48; DB 2; Length 30;
Best Local Similarity 45.0%; Pred. No. 8.5e+02;
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 34 REPTSSPLYSTYTDGNDTVT 53
DB 5 RNPTVPTDYDVTDRSRLT 24

RESULT 3
Q7M1G3 PRELIMINARY; PRT; 46 AA.
AC Q7M1G3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Arabinogalactan protein (Fragments).
OS Lolium multiflorum (Italian ryegrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Poace; Lolium.
OX NCBI_TaxID=4521;
RN [1]
RP SEQUENCE.
RX MEDLINE=90147544; PubMed=2695069;
RA Gleeson P.A., McNamara M., Wattenhall R.E.H., Stone B.A.,
RA Fincher G.B.;
RT "Characterization of the hydroxyproline-rich protein core of an
RT arabinogalactan-protein secreted from suspension-cultured Lolium
RT multiflorum (Italian ryegrass) endosperm cells.";
RL Biochem. J. 264:857-862(1989).
DR PIR; S07073; S07073.
FT NON_TER 1
FT NON_TER 46
SQ SEQUENCE 46 AA; 4173 MW; EF61188B0C8A08D0 CRC64;

Query Match 11.0%; Score 47.5; DB 2; Length 46;
Best Local Similarity 47.8%; Pred. No. 1.6e+03;
Matches 11; Conservative 2; Mismatches 7; Indels 3; Gaps 1;

QY 2 PLRQNHPEVA---TSPSSPPA 21
DB 18 FVPEASTAPVAAPTXXSPSPA 40

RESULT 4
Q8YFV1 PRELIMINARY; PRT; 47 AA.
AC Q8YFV1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein BME11419.
GN OrderedLocustNames=BME11419.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;

```

```

[1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA Salkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Hasekorn R., Kyripides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
DR EMBL; AE009579; AAL52600.1; -.
DR PIR; AE3429; AE3429.
KW Complete proteome.
SQ SEQUENCE 47 AA; 5180 MW; 989EE40500EC57D0 CRC64;

Query Match 11.0%; Score 47.5; DB 2; Length 47;
Best Local Similarity 45.9%; Pred. No. 1.6e+03;
Matches 17; Conservative 3; Mismatches 12; Indels 5; Gaps 3;

QY 4 PRONHEPVATSP--SSQPAETHP-TTLQGAIRREPT 37
DB 12 PRP--DAVLRSPDSDAPQPAATAPNTTQNRWRELKT 46

RESULT 5
Q86LU3 PRELIMINARY; PRT; 43 AA.
AC Q86LU3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=870;
RX MEDLINE=22511025; PubMed=12624019;
RA Leon F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY176656; AAO19461.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4750 MW; 1D0A81C40B323D35 CRC64;

Query Match 10.9%; Score 47; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 1.6e+03;
Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 44 YTTDGNDFVTSSDGL---WNNNTQLFLEHSLTLTA 76
DB 3 FVTDLSDLSNTADGLGQAW-YKISQVAIEHILTA 37

RESULT 6
Q81OR6 PRELIMINARY; PRT; 43 AA.
AC Q81OR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.

```

```

OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1005, 1034, 1283, and 340;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY135729; AAN62570.1; -
DR EMBL; AY135732; AAN62573.1; -
DR EMBL; AY135734; AAN62575.1; -
DR EMBL; AY135735; AAN62576.1; -
DR EMBL; AY135738; AAN62579.1; -
DR EMBL; AY176655; AA019460.1; -
DR EMBL; AY176657; AA019462.1; -
DR EMBL; AY176658; AA019463.1; -
DR EMBL; AY176660; AA019465.1; -
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4736 MW; 1D0A81C40B2A9E35 CRC64;

Query Match 10.6%; Score 46; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 2e+03;
Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 44 YTTDGNVTVESSDGL---WNNNOTQLFLEHSLTLTA 76
Db 3 FVTDLSSDLNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 7
Q810R7
ID Q810R7 PRELIMINARY; PRT; 43 AA.
AC Q810R7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T89, and 1962;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY135730; AAN62571.1; -
DR EMBL; AY176652; AA019457.1; -
DR EMBL; AY135733; AAN62574.1; -
DR EMBL; AY176654; AA019459.1; -
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4766 MW; 1D0A81C40B2A8E34 CRC64;

Query Match 10.6%; Score 46; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 2e+03;
Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 44 YTTDGNVTVESSDGL---WNNNOTQLFLEHSLTLTA 76
Db 3 FVTDLSSDLNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 8
Q810R7
ID Q810R7 PRELIMINARY; PRT; 43 AA.
AC Q810R7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=T89, and 1962;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY135730; AAN62571.1; -
DR EMBL; AY176652; AA019457.1; -
DR EMBL; AY135733; AAN62574.1; -
DR EMBL; AY176654; AA019459.1; -
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4766 MW; 1D0A81C40B2A8E34 CRC64;

Query Match 10.6%; Score 46; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 2e+03;
Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 44 YTTDGNVTVESSDGL---WNNNOTQLFLEHSLTLTA 76
Db 3 FVTDLSSDLNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 9
Q829S4
ID Q829S4 PRELIMINARY; PRT; 49 AA.
AC Q829S4;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF 5.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid pJTPS1.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu R., Akaishi K., Negishi H., Tanaka H., Ichinose Y.,
RA Shiraishi T., Yamada T.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB015669; BAA32212.1; -
KW Plasmid.
SQ SEQUENCE 49 AA; 5166 MW; A416D6945E2E8241 CRC64;

Query Match 10.6%; Score 46; DB 2; Length 49;
Best Local Similarity 36.8%; Pred. No. 2.4e+03;
Matches 14; Conservative 5; Mismatches 13; Indels 6; Gaps 2;

QY 4 PRONHEPVA--TSPSPQPAETHPTTLQGAIRREPTSS 39
Db 12 PAQSQKSIASITAPFSSPPAPRPPPSRQ----RRPTPS 45

RESULT 10
Q7UVZ0
ID Q7UVZ0 PRELIMINARY; PRT; 50 AA.
AC Q7UVZ0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
```

```

Q81SB4
ID Q81SB4 PRELIMINARY; PRT; 43 AA.
AC Q81SB4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY135737; AAN62578.1; -
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4766 MW; 1D0A81C41F7BDB35 CRC64;

Query Match 10.6%; Score 46; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 2e+03;
Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 44 YTTDGNVTVESSDGL---WNNNOTQLFLEHSLTLTA 76
Db 3 FVTDLSSDLNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 9
Q829S4
ID Q829S4 PRELIMINARY; PRT; 49 AA.
AC Q829S4;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ORF 5.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid pJTPS1.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu R., Akaishi K., Negishi H., Tanaka H., Ichinose Y.,
RA Shiraishi T., Yamada T.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB015669; BAA32212.1; -
KW Plasmid.
SQ SEQUENCE 49 AA; 5166 MW; A416D6945E2E8241 CRC64;

Query Match 10.6%; Score 46; DB 2; Length 49;
Best Local Similarity 36.8%; Pred. No. 2.4e+03;
Matches 14; Conservative 5; Mismatches 13; Indels 6; Gaps 2;

QY 4 PRONHEPVA--TSPSPQPAETHPTTLQGAIRREPTSS 39
Db 12 PAQSQKSIASITAPFSSPPAPRPPPSRQ----RRPTPS 45

RESULT 10
Q7UVZ0
ID Q7UVZ0 PRELIMINARY; PRT; 50 AA.
AC Q7UVZ0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
```


FT NON_TER 43 43
SQ SEQUENCE 43 AA; 4770 MW; 130461240B2A9E35 CRC64;
Query Match 10.2%; Score 44; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 3.2e+03;
Matches 12; Conservative 9; Mismatches 11; Indels 4; Gaps 2;
QY 44 YTTDGNQDTVTESDGL---WNNNQTLFLEHSLTLTA 76
: || : : : || | : : : || : ||
Db 3 FVTDLSSDFSNADGLQAW-YKISQVAVEHILTA 37

Search completed: June 29, 2005, 09:46:13
Job time : 49.8546 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:09:20 ; Search time 25.332 Seconds
(without alignments)
849.017 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225

Sequence: 1 PLYSTTGNDVTSSDGL.....NNQTQLFLEHLLTANTTKG 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 68540

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	20.9	43	2 Q86LU3	Q86LU3 cryptospori
2	46	20.4	43	2 Q81OR6	Q81OR6 cryptospori
3	46	20.4	43	2 Q81OR7	Q81OR7 cryptospori
4	46	20.4	43	2 Q81SB4	Q81SB4 cryptospori
5	45.5	20.2	48	2 Q6WV04	Q6WV04 saccharomyc
6	44	19.6	43	2 Q86LU1	Q86LU1 cryptospori
7	44	19.6	43	2 Q86LU6	Q86LU6 cryptospori
8	44	19.6	46	2 Q8J4I8	Q8J4I8 chimpanzee
9	44	19.6	49	2 Q8J4I5	Q8J4I5 chimpanzee
10	43.5	19.3	41	2 Q9J7T4	Q9J7T4 mus musculu
11	43	19.1	43	2 Q86LU2	Q86LU2 cryptospori
12	43	19.1	43	2 Q86LU5	Q86LU5 cryptospori
13	42.5	18.9	46	2 Q88390	Q88390 simian t-ly
14	42	18.7	43	2 Q86LU4	Q86LU4 cryptospori
15	42	18.7	46	2 Q7KYS6	Q7KYS6 homo sapien
16	42	18.7	47	2 Q8J4I9	Q8J4I9 chimpanzee
17	41.5	18.4	47	2 Q75JX4	Q75JX4 dictyosteli
18	41	18.2	29	2 Q86233	Q86233 human rotav
19	41	18.2	40	2 Q83DR8	Q83DR8 coxiella bu
20	41	18.2	41	2 Q11552	Q11552 human immu
21	40.5	18.0	49	2 Q91FZ1	Q91FZ1 chilo iride
22	40	17.8	49	2 Q7XA54	Q7XA54 gnetum gnem
23	39.5	17.6	46	2 Q82209	Q82209 human t-lym
24	39	17.3	28	2 Q998S5	Q998S5 turkey ente
25	39	17.3	30	2 Q9KT75	Q9KT75 vibrio chol
26	39	17.3	38	2 Q7RD73	Q7RD73 plasmodium
27	39	17.3	47	2 Q65SH0	Q65SH0 manheimia
28	39	17.3	50	2 Q7SEZ1	Q7SEZ1 neurospora
29	38.5	17.1	31	2 Q78501	Q78501 human immu
30	38.5	17.1	41	2 Q9R5W5	Q9R5W5 neisseria g
31	38	16.9	29	2 Q86237	Q86237 human rotav

32 38 16.9 29 2 Q86238 human rotav
33 38 16.9 29 2 Q86239 human rotav
34 38 16.9 38 2 Q71V48 human sapien
35 38 16.9 46 2 Q8UYK0 wheat dwarf
36 38 16.9 46 2 Q8UYK2 wheat dwarf
37 38 16.9 46 2 Q8UYK4 wheat dwarf
38 38 16.9 46 2 Q7SZJ6 accipiter n
39 38 16.9 46 2 Q7SZK0 Q7SZK0 aquila chry
40 38 16.9 46 2 Q7SZK2 Q7SZK2 circus spil
41 38 16.9 46 2 Q7SZK4 Q7SZK4 accipiter g
42 38 16.9 46 2 Q7SZK6 Q7SZK6 milvus migr
43 38 16.9 46 2 Q8J4I7 chimpanzee
44 38 16.9 46 2 Q9JNX5 human immu
45 38 16.9 50 2 Q7V009 Q7V009 prochloroco

ALIGNMENTS

RESULT 1
Q86LU3 PRELIMINARY; PRT; 43 AA.
ID Q86LU3
AC Q86LU3;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=870;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from humans and animals by using a heteroduplex mobility assay and nucleic acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AV176656; AAO19461.1; -;
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4750 MW; 1D0A81C40B323D35 CRC64;
Query Match 20.9%; Score 47; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 1.3e+02;
Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;
QY 5 YTTDGNVTSSDGL---WNNQTQLFLEHLLTA 37
DB 3 FVTDLSLNTAGLGGQAW-YKISQVAIEHLLTA 37
RESULT 2
Q81OR6 PRELIMINARY; PRT; 43 AA.
ID Q81OR6
AC Q81OR6;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1005, 1034, 1283, and 340;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from

RT humans and animals by using a heteroduplex mobility assay and nucleic
 RT acid sequencing based on a small double-stranded RNA element.";

RL J. Clin. Microbiol. 41:981-992(2003).

DR EMBL; AY135729; AAN62570.1; -

DR EMBL; AY135732; AAN62573.1; -

DR EMBL; AY135734; AAN62575.1; -

DR EMBL; AY135735; AAN62576.1; -

DR EMBL; AY135738; AAN62579.1; -

DR EMBL; AY135738; AAN62579.1; -

DR EMBL; AY176655; AAO19460.1; -

DR EMBL; AY176657; AAO19462.1; -

DR EMBL; AY176658; AAO19463.1; -

DR EMBL; AY176660; AAO19465.1; -

KW Hypothetical protein.

FT NON_TER 1

FT NON_TER 43

SQ SEQUENCE 43 AA; 4736 MW; 1D0A81C40B2A9E35 CRC64;

Query Match 20.4%; Score 46; DB 2; Length 43;

Best Local Similarity 33.3%; Pred. No. 1.7e+02;

Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 5 YTTDGNDFVTSSDGL---WNNQTLFLEHLLTA 37

:||: : : :||| | :||: :||: |||

Db 3 FVTDLSSDLNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 3

Q810R7

ID Q810R7 PRELIMINARY; PRT; 43 AA.

AC Q810R7;

DT 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)

DE Hypothetical protein (Fragment).

OS Cryptosporidium parvum.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

OC Cryptosporidiidae; Cryptosporidium.

OX NCBI_TaxID=5807;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=T89, and 1962;

RX MEDLINE=22511025; PubMed=12624019;

RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;

RT "Molecular epidemiological analysis of Cryptosporidium isolates from

RT humans and animals by using a heteroduplex mobility assay and nucleic

RT acid sequencing based on a small double-stranded RNA element.";

RL J. Clin. Microbiol. 41:981-992(2003).

DR EMBL; AY135730; AAN62571.1; -

DR EMBL; AY176652; AAO19457.1; -

DR EMBL; AY135733; AAN62574.1; -

DR EMBL; AY176654; AAO19459.1; -

KW Hypothetical protein.

FT NON_TER 1

FT NON_TER 43

SQ SEQUENCE 43 AA; 4766 MW; 1D0A81C40B2A8E34 CRC64;

Query Match 20.4%; Score 46; DB 2; Length 43;

Best Local Similarity 33.3%; Pred. No. 1.7e+02;

Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 5 YTTDGNDFVTSSDGL---WNNQTLFLEHLLTA 37

:||: : : :||| | :||: :||: |||

Db 3 FVTDLSSDLNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 4

Q81SB4

ID Q81SB4 PRELIMINARY; PRT; 43 AA.

AC Q81SB4;

DT 01-MAR-2003 (TReMBLrel. 23, Created)

DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)

DE Hypothetical protein (Fragment).

OS Cryptosporidium parvum.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

OC Cryptosporidiidae; Cryptosporidium.

OX NCBI_TaxID=5807;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22511025; PubMed=12624019;

RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;

RT "Molecular epidemiological analysis of Cryptosporidium isolates from

RT humans and animals by using a heteroduplex mobility assay and nucleic

RT acid sequencing based on a small double-stranded RNA element.";

RL J. Clin. Microbiol. 41:981-992(2003).

DR EMBL; AY135737; AAN62578.1; -

KW Hypothetical protein.

FT NON_TER 1

FT NON_TER 43

SQ SEQUENCE 43 AA; 4766 MW; 1D0A81C41F7DB35 CRC64;

Query Match 20.4%; Score 46; DB 2; Length 43;

Best Local Similarity 33.3%; Pred. No. 1.7e+02;

Matches 12; Conservative 10; Mismatches 10; Indels 4; Gaps 2;

QY 5 YTTDGNDFVTSSDGL---WNNQTLFLEHLLTA 37

:||: : : :||| | :||: :||: |||

Db 3 FVTDLSSDLNTDGLGQAW-YKISQVAVEHILTA 37

RESULT 5

Q6WV04

ID Q6WV04 PRELIMINARY; PRT; 48 AA.

AC Q6WV04;

DT 05-JUL-2004 (TReMBLrel. 27, Created)

DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)

DE Defective Sdc25p (Fragment).

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RA Brachet S., Dietrich F.S., Voegeli S., Zhang Z., Stuart L.T.,

RA Lerch A., Gaffney T., Philippsen P.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;

RA Zhang Z., Stuart L.T., Dietrich F.S.;

RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AY268135; AAR26281.1; -

FT NON_TER 1

FT NON_TER 48

SQ SEQUENCE 48 AA; 5443 MW; EE70B5821E357BFA CRC64;

Query Match 20.2%; Score 45.5; DB 2; Length 48;

Best Local Similarity 25.5%; Pred. No. 2.2e+02;

Matches 12; Conservative 9; Mismatches 9; Indels 17; Gaps 2;

QY 13 VTSSDGLW-----NNQTLFLE-----HSLLTANTTKG 42

:||: : : :||| | :||: :||: |||

Db 1 LTKSGNGWGDGVLIRHSANNNNNSLIDRGWFPFPPSYTVHSKRTTRG 47

RESULT 6

Q86LU1

ID Q86LU1 PRELIMINARY; PRT; 43 AA.

AC Q86LU1;

DT 01-JUN-2003 (TReMBLrel. 24, Created)

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)

DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)

DE Hypothetical protein (Fragment).

OS Cryptosporidium parvum.

OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;

```
OC Cryptosporidiidae; Cryptosporidium.
RN NCBI_TaxID=5807;
RP SEQUENCE FROM N.A.
RC STRAIN=1876;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY176661; AAO19466.1; --
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 43 43
SQ SEQUENCE 43 AA; 4770 MW; 130461240B2A9E35 CRC64;

Query Match 19.6%; Score 44; DB 2; Length 43;
Best Local Similarity 33.3%; Pred. No. 3e+02;
Matches 12; Conservative 9; Mismatches 11; Indels 4; Gaps 2;

QY 5 YTTDGNVTSSDGL---WNNQTLFLEHSLTA 37
   : : : : : : : : : : : : : : : : : :
DB 3 FVTDLSSDFSNTADGLGQAW-YKISQVAVEHILTA 37

RESULT 7
Q86LU6 PRELIMINARY; PRT; 43 AA.
AC Q86LU6;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium meleagridis
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=93969;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1689;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY176647; AAO19452.1; --
KW Hypothetical protein.
FT NON_TER 1 1
FT NON_TER 43 43
SQ SEQUENCE 43 AA; 4812 MW; FD31E1C40B323830 CRC64;

Query Match 19.6%; Score 44; DB 2; Length 43;
Best Local Similarity 29.7%; Pred. No. 3e+02;
Matches 11; Conservative 11; Mismatches 9; Indels 6; Gaps 2;

QY 5 YTTDGNVTSSDGL---WNNQTLFLEHSLTA 37
   : : : : : : : : : : : : : : : : : :
DB 3 FVTDLSSDLANTADGLGQAWYKISQTE--IEHILMTA 37

RESULT 8
Q8J4I8 PRELIMINARY; PRT; 46 AA.
AC Q8J4I8;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE Envelope glycoprotein (Fragment).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
```

```
RP SEQUENCE FROM N.A.
RA Trichel A.M., Rajakumar P.A., Murphey-Corb M.A.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AY118209; AAM78569.1; --
DR GO; GO:0019031; C:Viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 46 46
SQ SEQUENCE 46 AA; 4928 MW; 9A2C139D4777658E CRC64;

Query Match 19.6%; Score 44; DB 2; Length 46;
Best Local Similarity 48.0%; Pred. No. 3.3e+02;
Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 4 SYTTDGNVTVTSSDGLMNNNOTQL 28
   : : : : : : : : : : : : : : : : : :
DB 18 STTTPKADVNETSSCVKNNCTGL 42

RESULT 9
Q8J4I5 PRELIMINARY; PRT; 49 AA.
AC Q8J4I5;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DE Envelope glycoprotein (Fragment).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA Trichel A.M., Rajakumar P.A., Murphey-Corb M.A.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AY118212; AAM78572.1; --
KW Envelope protein.
FT NON_TER 1 1
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 5101 MW; 86D35EEB15D4FC94 CRC64;

Query Match 19.6%; Score 44; DB 2; Length 49;
Best Local Similarity 48.0%; Pred. No. 3.5e+02;
Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 4 SYTTDGNVTVTSSDGLMNNNOTQL 28
   : : : : : : : : : : : : : : : : : :
DB 21 STTTPKADVNETSSCVKNNCTGL 45

RESULT 10
Q99JT4 PRELIMINARY; PRT; 41 AA.
AC Q99JT4;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Basa S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Roha S.A., McSwain J.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettenan M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleka U., Smalilus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."; [2]
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
 RL Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005705; AAH05705.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 41 AA; 4720 MW; 7692D458655D3B05 CRC64;

Query Match 19.3%; Score 43.5; DB 2; Length 41;
Best Local Similarity 31.6%; Pred. No. 3.3e+02;
Matches 12; Conservative 3; Mismatches 18; Indels 5; Gaps 1;

Qy 2 LYSYTTDGDNDVTESDGLNANNQQLFLEHSLITANT 39
|| || | : ||| : | | | |
Db 3 LYLLRDANHTIL-----LWSKNYKCLGTVNHGLKVARS 35

RESULT	11
Q86LU2	
ID	Q86LU2 PRELIMINARY; PRT; 43 AA.
AC	Q86LU2;
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Hypothetical protein (Fragment).
OS	Cryptosporidium parvum.
OC	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC	Cryptosporidiidae; Cryptosporidium.
OX	NCBI_TaxID=5807;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=1059;
RX	MEDLINE=22511025; PubMed=12624019;
RA	Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT	"Molecular epidemiological analysis of Cryptosporidium
RT	humans and animals by using a heteroduplex mobility assay
RL	acid sequencing based on a small double-stranded RNA eluted
RL	J. Clin. Microbiol. 41:981-992(2003).
DR	EMBL; AV176659; AAO19464.1; -.
KW	Hypothetical protein.
FT	NON_TER 1
FT	NON_TER 43
SO	SEQUENCE 43 AA; 4754 MW; 249AA01240BA9E26 CRC64;

Query Match	19.1%	Score 43;	DB 2;	Length 43;
Best Local Similarity	33.3%	Pred. No. 4.1e+02;		
Matches 12; Conservative	8;	Mismatches 12;	Indels 4;	Gaps 2;

Qy	5	YTTDGN	DTVESS	DGI---	WNNTQ	QLFLEH	SLTA	37
Dp	3	FVTDL	SSDFANT	ADG	GOAV-	YKIS	OVAVEH	ILTA 37

RESULT 12	
Q86LU5	
ID Q86LU5	PRELIMINARY;
AC Q86LU5;	PRT;
DT 01-JUN-2003	43 AA.
(TREMBLrel. 24. Created)	

DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)	
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)	
DE	Hypothetical protein (Fragment).	
OS	Cryptosporidium felis.	
OC	Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida;	
OC	Cryptosporidiidae; Cryptosporidium.	
OX	NCBI_Taxid=83540;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=355;	
RX	MEDLINE=22511025; PubMed=12624019;	
RA	Leoni F., Gallimore C.I., Green J., McLauchlin J.;	
RT	"Molecular epidemiological analysis of Cryptosporidium isolates from	
RT	humans and animals by using a heteroduplex mobility assay and nucleic	
RT	acid sequencing based on a small double-stranded RNA element.";	
RL	J. Clin. Microbiol. 41:981-992(2003).	
DR	EMBL; AY176649; AAO19454.1; -.	
KW	Hypothetical protein.	
FT	NON TER 1	
FT	NON TER 43	
FT	NON TER 43	
SO	SEQUENCE 43 AA; 4812 MW; 1C4F81C40B2A9B21 CRC64;	

Query Match 19.1%; Score 43; DB 2; Length 43;
Best Local Similarity 30.6%; Pred. No. 4.1e+02;
Matches 11; Conservative 11; Mismatches 10; Indels

Qy 5 YTTDGDNTVTSSDGL---WNNNQQLFLEHSLLLTA 37
: || : : : || | : : : || : ||
Dp 3 FVTLSSDLSNTADLGGAW-YKISOVEVPHIIMTA 37

[illegible]

Query Match 18.9%; Score 42.5; DB 2; Length 46;
Best Local Similarity 39.4%; Pred. No. 5.1e+02;
Matches 13; Conservative 3; Mismatches 14; Indels 3; Gaps 2;

QY 2 LYSYTTDGDNDVTVES--SDGLWNNNQTLFLEH 32
||| ||| | : : ||| | : | |
Db 13 LYKYFTDKDPLPMDNALSGLWTFNHLNV-LTH 44

RESULT 14	
Q86IU4	
ID	Q86IU4
AC	Q86IU4;
DT	01-JUN-2003 (TREMELrel. 24, Created)
	PRELIMINARY;
	PRT;
	43 AA.

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein (Fragment).
OS Cryptosporidium felis.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=83540;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1626;
RX MEDLINE=22511025; PubMed=12624019;
RA Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT "Molecular epidemiological analysis of Cryptosporidium isolates from
RT humans and animals by using a heteroduplex mobility assay and nucleic
RT acid sequencing based on a small double-stranded RNA element.";
RL J. Clin. Microbiol. 41:981-992(2003).
DR EMBL; AY176651; AAC19456.1; --
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 43
SQ SEQUENCE 43 AA; 4798 MW; CB1F81C40B323833 CRC64;

Query Match 18.7%; Score 42; DB 2; Length 43;
Best Local Similarity 30.6%; Pred. No. 5.5e+02;
Matches 11; Conservative 10; Mismatches 11; Indels 4; Gaps 2;

QY 5 YTDGNDVTVESDGL---WNNNQTLFLEHSLTLTA 37
: || : : : ||| | : || : |||
DB 3 FVTDLSLNTADGLGQAW-YKISQAEIHIIMTA 37
: || : : : ||| | : || : |||

RESULT 15
Q7KYS6
ID Q7KYS6 PRELIMINARY; PRT; 46 AA.
AC Q7KYS6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Keen T.J.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U85997; AAB46606.1; --
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 46 AA; 5123 MW; 64D4B2844B295426 CRC64;

Query Match 18.7%; Score 42; DB 2; Length 46;
Best Local Similarity 34.3%; Pred. No. 5.9e+02;
Matches 12; Conservative 5; Mismatches 16; Indels 2; Gaps 1;

QY 9 GNDVTVES--SDGLMNNQTLFLEHSLTLTATTK 41
: || : : : ||| | : || : |||
DB 1 GCTTPESDLEERSVEQDSTELFTNHRHLTAETPR 35
: || : : : ||| | : || : |||

Search completed: June 29, 2005, 09:46:14
Job time : 26.332 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:33:11 ; Search time 7.59136 Seconds
(without alignments)
413.004 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81
Perfect score: 225
Sequence: 1 PLYSYTGDNDVTSSDGL.....NNQTQLFLEHSLTANTTKG 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 272444

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCUTS_COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfilesl.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	21.3	30	4	US-09-282-029A-470
2	48	21.3	30	4	US-09-282-029A-471
3	48	21.3	30	4	US-09-434-355A-470
4	48	21.3	30	4	US-09-434-355A-471
5	45.5	20.2	40	1	US-08-188-228-22
6	45.5	20.2	40	1	US-08-332-643-22
7	45.5	20.2	40	1	US-08-332-638-22
8	44.5	19.8	31	1	US-08-190-802A-244
9	44.5	19.8	31	3	US-08-477-346-244
10	44.5	19.8	31	3	US-08-477-346-244
11	44.5	19.8	31	4	US-08-487-072A-244
12	44	19.6	20	1	US-08-218-025A-65
13	42	18.7	40	1	US-08-188-228-16
14	42	18.7	40	1	US-08-332-643-16
15	42	18.7	40	1	US-08-332-638-16
16	41.5	18.4	32	4	US-08-708-620A-4
17	41	18.2	42	4	US-09-434-355A-529
18	40	17.8	14	4	US-09-570-921-141
19	40	17.8	42	4	US-09-282-029A-6
20	40	17.8	42	4	US-09-185-908-6
21	40	17.8	42	4	US-09-434-355A-6
22	40	17.8	49	3	US-09-217-228-9
23	39.5	17.6	30	1	US-08-190-802A-123
24	39.5	17.6	30	3	US-08-477-346-123
25	39.5	17.6	30	3	US-08-473-089-123
26	39.5	17.6	30	4	US-08-487-072A-123
27	39	17.3	40	1	US-08-188-228-20

28 39 17.3 40 1 US-08-332-643-20 Sequence 20, Appl
29 39 17.3 40 1 US-08-332-638-20 Sequence 20, Appl
30 39 17.3 41 2 US-08-773-008-5 Sequence 5, Appli
31 38.5 17.1 30 1 US-08-190-802A-162 Sequence 162, App
32 38.5 17.1 30 3 US-08-477-346-162 Sequence 162, App
33 38.5 17.1 30 3 US-08-473-089-162 Sequence 162, App
34 38.5 17.1 30 4 US-08-487-072A-162 Sequence 162, App
35 38.5 17.1 31 1 US-08-190-802A-236 Sequence 236, App
36 38.5 17.1 31 3 US-08-477-346-236 Sequence 236, App
37 38.5 17.1 31 3 US-08-473-089-236 Sequence 236, App
38 38.5 17.1 31 4 US-08-487-072A-236 Sequence 236, App
39 38 16.9 21 3 US-09-252-586-18 Sequence 18, Appl
40 38 16.9 31 4 US-08-437-943D-76 Sequence 76, Appl
41 38 16.9 35 2 US-08-749-526-5 Sequence 5, Appli
42 37.5 16.7 46 4 US-08-875-553D-34 Sequence 34, Appl
43 37.5 16.7 46 4 US-08-875-553D-35 Sequence 35, Appl
44 37.5 16.7 47 3 US-08-331-625A-29 Sequence 29, Appl
45 37.5 16.7 47 3 US-09-494-151-29 Sequence 29, Appl

ALIGNMENTS

RESULT 1
US-09-282-029A-470
; Sequence 470, Application US/09282029A
; Patent No. 6723700
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/282, 029A
; CURRENT FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 480
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 470
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-282-029A-470

Query Match 21.3%; Score 48; DB 4; Length 30;
Best Local Similarity 45.8%; Pred. No. 10;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 2;

QY 20 LWNNTQLFLEHSLTANTTKG 42
DB 2 MW---QVTAFLDHNIVTAQTWK 22

RESULT 2
US-09-282-029A-471
; Sequence 471, Application US/09282029A
; Patent No. 6723700
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: Gour, Barbara J.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C1
; CURRENT APPLICATION NUMBER: US/09/282, 029A
; CURRENT FILING DATE: 1999-03-30
; NUMBER OF SEQ ID NOS: 480
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 471
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-282-029A-471

Query Match 21.3%; Score 48; DB 4; Length 30;

```
Best Local Similarity 45.8%; Pred. No. 10;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 2;

QY 20 LMNNQTLFLEHSLLTANTT-KG 42
DB 2 MW---QVTAFLDHNIVTAQTWK 22

RESULT 3
US-09-434-355A-470
; Sequence 470, Application US/09434355A
; Patent No. 6830894
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355A
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 470
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-434-355A-470

Query Match 21.3%; Score 48; DB 4; Length 30;
Best Local Similarity 45.8%; Pred. No. 10;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 2;

QY 20 LMNNQTLFLEHSLLTANTT-KG 42
DB 2 MW---QVTAFLDHNIVTAQTWK 22

RESULT 4
US-09-434-355A-471
; Sequence 471, Application US/09434355A
; Patent No. 6830894
; GENERAL INFORMATION:
; APPLICANT: Blaschuk, Orest W.
; APPLICANT: James Mathew Symonds
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR MODULATING CLAUDIN-MEDIATED
; FILE REFERENCE: 100086.409C2
; CURRENT APPLICATION NUMBER: US/09/434,355A
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 471
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-434-355A-471

Query Match 21.3%; Score 48; DB 4; Length 30;
Best Local Similarity 45.8%; Pred. No. 10;
Matches 11; Conservative 5; Mismatches 4; Indels 4; Gaps 2;

QY 20 LMNNQTLFLEHSLLTANTT-KG 42
DB 2 MW---QVTAFLDHNIVTAQTWK 22

RESULT 5
US-08-188-228-22
; Sequence 22, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,228
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,460
; FILING DATE: 19 APR 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 APR 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5597725and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-188-228-22

Query Match 20.2%; Score 45.5; DB 1; Length 40;
Best Local Similarity 31.7%; Pred. No. 32;
Matches 13; Conservative 5; Mismatches 8; Indels 15; Gaps 1;

QY 2 LYSVTTDNDVTSSDGLMNNQTLFLEHSLLTANTTKG 42
DB 2 LATVAYEGNDSVAESL-----SSLESGETEG 27

RESULT 6
US-08-332-643-22
; Sequence 22, Application US/08332643
; Patent No. 5639634
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```



```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,643
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,643
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5639634and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; US-08-332-643-22

Query Match          20.2%; Score 45.5; DB 1; Length 40;
Best Local Similarity 31.7%; Pred. No. 32;
Matches 13; Conservative 5; Mismatches 8; Indels 15; Gaps 1;

QY  2 LYSYTTDGDNTVTSSDGLWNNNQTQLFLEHSLLTANTTKG 42
   | : | | | | | | | | | | | | | | | | | | |
Db  2 LATYAYEGNDSVAESL-----SSLESQTTEG 27

RESULT 7
US-08-332-638-22
; Sequence 22, Application US/08332638
; Patent No. 5646250
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,638
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 APR 1992
; APPLICATION NUMBER: US/08/049,460
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5646250and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-332-638-22

Query Match          20.2%; Score 45.5; DB 1; Length 40;
Best Local Similarity 31.7%; Pred. No. 32;
Matches 13; Conservative 5; Mismatches 8; Indels 15; Gaps 1;

QY  2 LYSYTTDGDNTVTSSDGLWNNNQTQLFLEHSLLTANTTKG 42
   | : | | | | | | | | | | | | | | | | | | |
Db  2 LATYAYEGNDSVAESL-----SSLESQTTEG 27

RESULT 8
US-08-190-802A-244
; Sequence 244, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: YCU7 r1, Fig. 48
; US-08-190-802A-244

Query Match          19.8%; Score 44.5; DB 1; Length 31;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY  4 SYTTDGDNTVTSSDQ---LWN 22
   | : | | | | | | | | | |
Db  10 AYSPDGSRVVTASDQKIKVWD 31

RESULT 9
US-08-477-346-244
; Sequence 244, Application US/08477346
; Patent No. 6262023
; GENERAL INFORMATION:
```

APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
ZIP: 20006-1812
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,346
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,072
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: YCU7 rI, Fig. 48
US-08-477-346-244

Query Match 19.8%; Score 44.5; DB 3; Length 31;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 4 SYTTDGNNTVTSSDG---LWN 22
DB 10 AYSPDGSRVVTASEDGKIKWVD 31

RESULT 10
US-08-473-089-244
Sequence 244, Application US/08473089
Patent No 6342368
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
ZIP: 20006-1812
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,089
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
INDIVIDUAL ISOLATE: YCU7 rI, Fig. 48
US-08-473-089-244

Query Match 19.8%; Score 44.5; DB 3; Length 31;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 4 SYTTDGNNTVTSSDG---LWN 22
DB 10 AYSPDGSRVVTASEDGKIKWVD 31

RESULT 11
US-08-487-072A-244
Sequence 244, Application US/08487072A
Patent No. 6423684
GENERAL INFORMATION:
APPLICANT: Mochly-Rosen, Daria
APPLICANT: Ron, Dorit
TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
ZIP: 20006-1812
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,072A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: MURASHIGE, KATE H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 2550-0025.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
INFORMATION FOR SEQ ID NO: 244:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO

```
; ANTI-SENSE: NO
; ORIGINAL SOURCE: YCU7 r1, Fig. 48
; INDIVIDUAL ISOLATE: YCU7 r1, Fig. 48
US-08-487-072A-244

Query Match 19.8%; Score 44.5; DB 4; Length 31;
Best Local Similarity 40.9%; Pred. No. 31;
Matches 9; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 4 SYTTDGNVTVTSSDG---LWN 22
   ||: ||: ||: ||: ||: ||:
Db 10 AVSPDGRVVTASEDKIKVWD 31

RESULT 12
US-08-218-025A-65
; Sequence 65, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Welner, David B.
; APPLICANT: Ugen, Kenneth E.
; APPLICANT: Williams, William V.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; TITLE OF INVENTION: and Treating Certain HIV Infected Patients
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 5556744tistown Road
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,025A
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/891,451
; FILING DATE: 29-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST33A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-218-025A-65

Query Match 19.6%; Score 44; DB 1; Length 20;
Best Local Similarity 39.1%; Pred. No. 21;
Matches 9; Conservative 4; Mismatches 2; Indels 8; Gaps 1;

QY 1 PLYSYTTDGNVTVTSSDGLWNN 23
   ||: ||: ||: ||: ||: ||:
Db 2 PLFNSTWNGNT-----WNN 16

RESULT 13
US-08-188-228-16
; Sequence 16, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two First National Plaza, 20 South Clark
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,643
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,643
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5639634and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-332-643-16

Query Match      18.7%; Score 42; DB 1; Length 40;
Best Local Similarity 39.1%; Pred. No. 94;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      2 LYSYTTDGNQDVTSSDGLMNN 24
      | : | : | : | : | : | : | :
Db      2 LQYAFEGNGSVAESLSLSDNS 24
      | : | : | : | : | : | : | :

Search completed: June 29, 2005, 09:49:07
Job time : 8.59136 secs

;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-332-638-16

Query Match      18.7%; Score 42; DB 1; Length 40;
Best Local Similarity 39.1%; Pred. No. 94;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      2 LYSYTTDGNQDVTSSDGLMNN 24
      | : | : | : | : | : | : | :
Db      2 LQYAFEGNGSVAESLSLSDNS 24
      | : | : | : | : | : | : | :

Search completed: June 29, 2005, 09:49:07
Job time : 8.59136 secs

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,643
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,643
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5639634and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-332-643-16

Query Match      18.7%; Score 42; DB 1; Length 40;
Best Local Similarity 39.1%; Pred. No. 94;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY      2 LYSYTTDGNQDVTSSDGLMNN 24
      | : | : | : | : | : | : | :
Db      2 LQYAFEGNGSVAESLSLSDNS 24
      | : | : | : | : | : | : | :

RESULT 15
US-08-332-638-16
; Sequence 16, Application US/08332638
; Patent No. 5646250
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,638
; FILING DATE: 01-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 APR 1992
; APPLICATION NUMBER: US/08/049,460
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5646250and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:27:45 ; Search time 44.6208 Seconds
(without alignments)
720.211 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLHLADSVAG.....PFKKEVQQLRPHKSCIHQRE 334
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 11837

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	3.4	42	2 I70082	glycoprotein Ib al
2	56	3.2	43	2 JU0404	hypothetical 5.2K
3	55	3.1	40	2 B69408	hypothetical prote
4	53.5	3.1	46	2 S07073	arabinogalactan pr
5	50	2.9	34	2 D81044	hypothetical prote
6	49	2.8	25	2 I40692	cenA protein (iga1
7	49	2.8	42	2 T07030	extensin - tomato
8	49	2.8	45	2 H44530	T-cell receptor al
9	48.5	2.8	27	2 A90629	thr operon leader
10	48.5	2.8	27	2 A85480	thr operon leader
11	47.5	2.7	47	2 AE3429	hypothetical prote
12	47	2.7	50	2 I38970	fibroblast growth
13	46	2.6	46	2 F81142	hypothetical prote
14	45.5	2.6	50	2 A85902	hypothetical prote
15	45	2.6	35	2 S20766	Ig heavy chain V r
16	45	2.6	47	2 B81103	hypothetical prote
17	44.5	2.5	29	2 C61384	tracheal mucin gly
18	44.5	2.5	44	2 A11110	Ig heavy chain V r
19	44.5	2.5	50	2 AH2135	hypothetical prote
20	43.5	2.5	41	2 S38918	hypothetical prote
21	43.5	2.5	49	2 S49365	Ig mu chain transm
22	43	2.5	28	2 FH1363	Ig heavy chain DJ
23	43	2.5	31	2 T29628	hypothetical prote
24	43	2.5	32	2 S58524	alpha-complex prot
25	43	2.5	35	2 S30112	aminoglycoside N6'
26	43	2.5	41	2 D45495	beta-defensin-4 -
27	43	2.5	41	2 C30208	hypothetical prote
28	43	2.5	43	2 S49760	extensin - tomato
29	43	2.5	43	2 B42057	fibroblast growth

30	43	2.5	47	2 A41266	fibroblast growth
31	43	2.5	50	2 A47694	nitrate reductase
32	42.5	2.4	40	2 S33406	Ig heavy chain V r
33	42.5	2.4	43	2 A95035	hypothetical prote
34	42.5	2.4	46	2 A49281	pol protein - simi
35	42.5	2.4	49	2 S25433	neural cell adhesi
36	42	2.4	22	2 I41299	T-cell receptor be
37	42	2.4	28	2 I52394	fibronectin, splic
38	42	2.4	41	2 C87603	hypothetical prote
39	42	2.4	45	2 B46346	4.8K nonstructural
40	42	2.4	49	2 PL0096	Ig heavy chain V r
41	42	2.4	49	2 S68219	defensin AMP1 - Cl
42	42	2.4	50	2 B49410	t-complex polypept
43	41.5	2.4	36	2 A37172	collagen alpha 1(x
44	41.5	2.4	37	2 E97596	hypothetical prote
45	41.5	2.4	41	2 PQ0104	microbial serine p

ALIGNMENTS

RESULT 1

I70082 Glycoprotein Ib alpha variant B - human (fragment)

C:Species: Homo sapiens (man)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C:Accession: I70082
R:Lopez, J.A.; Ludwig, E.H.; McCarthy, B.J.
J. Biol. Chem. 267, 10055-10061, 1992
A:Title: Polymorphism of human glycoprotein Ib alpha results from a variable number of t
actions.

A:Reference number: I55355; MUID:92250564; PMID:1577776
A:Accession: I70082
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-42 <RES>
A:Cross-references: UNIPROT:Q16469; GB:S34439; NID:G249178; PIDN:AB22153.1; PID:G249179
C:Superfamily: platelet glycoprotein Ib alpha chain; leucine-rich alpha-2-glycoprotein r
C:Keywords: glycoprotein

Query Match 3.4%; Score 59; DB 2; Length 42;
Best Local Similarity 29.0%; Pred. No. 9.8e+02; Mismatches 15; Indels 22; Gaps 3;
Matches 18; Conservative 7;

QY	168	PRTTVPTMTVTSTVPTTSIPTTSVPVTTVTSTVTPMPPLRQNHVPVATSPSSP 227
DB	2	PTSEAP-----SPTTPTSTSEAPSPPT-----PEPTSEP-APSPPTP 39
QY	228	QP 229
DB	40	EP 41

RESULT 2

JU0404 Hypothetical 5.2K protein - clover yellow mosaic virus

C:Species: clover yellow mosaic virus
C>Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 09-Jul-2004
C:Accession: JU0404
R:Abouneidar, M.G.; Lai, R.
J. Gen. Virol. 70, 1871-1875, 1989
A:Title: Nucleotide sequence of the 3'-terminal region of clover yellow mosaic virus RNA
A:Reference number: JU0402; MUID:89293092; PMID:2738582
A:Accession: JU0404
A:Molecule type: genomic RNA
A:Residues: 1-43 <ABO>
A:Cross-references: UNIPROT:P16485; GB:D00485; NID:G221218; PIDN:BAA00374.1; PID:d100082

A>Note: this protein is encoded by sequence in front of the coat protein-encoding sequen
Query Match 3.2%; Score 56; DB 2; Length 43;
Best Local Similarity 40.0%; Pred. No. 1.5e+03; Mismatches 9; Indels 4; Gaps 1;
Matches 12; Conservative 5;

```

Qy 207 VPMPLPRQNHPEVATSPSPQPAETHPTT 236
   :  :  |||  :  |||  :  |||  :  |||
Db 14 LPRLSRPQSHLPAAQT-----PQRLSYPKT 39

RESULT 3
B69408
hypothetical protein AF1267 - Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Species: Archaeoglobus fulgidus
C:Accession: B69408
F:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
  ; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
  Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
  Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaea
  A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69408
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-40 <KLE>
A:Cross-references: UNIPROT:O29001; GB:AE001017; GB:AE000782; NID:G2689340; PIDN:AAB89999

Query Match 3.1%; Score 55; DB 2; Length 40;
Best Local Similarity 28.6%; Pred. No. 1.6e+03;
Matches 14; Conservative 3; Mismatches 16; Indels 16; Gaps 2;

Qy 17 SVAGSVKVGGEAGPSVTLPHYSGAVTSMCNRGSCSLFTQCGIVWTN 65
   :  :  |||  :  |||  :  |||  :  |||
Db 2 AVAALAMYGGTCGACAVLACN-----WNVREC-----GIIWKN 34

RESULT 4
S07073
arabinogalactan protein - Italian ryegrass (fragments)
C:Species: Lolium multiflorum (Italian ryegrass)
C:Date: 21-Nov-1993 #sequence_revision 01-Dec-1995 #text_change 09-Jul-2004
C:Accession: S07073
F:Gleeson, P.A.; McNamara, M.; Wattenhall, R.E.H.; Stone, B.A.; Fincher, G.B.
Biochem. J. 264, 857-862, 1989
A:Title: Characterization of the hydroxyproline-rich protein core of an arabinogalactan-
  A:Reference number: S07073; MUID:90147544; PMID:2695069
A:Accession: S07073
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-46 <GLE>
A:Cross-references: UNIPROT:Q7M1G3
A:Note: 19-His and 23-Leu were also found
C:Keywords: hydroxyproline
F:6,7,9,11,14,15,16,20,26,30,34,36,39/Modified site: hydroxyproline (Pro) #status experi

Query Match 3.1%; Score 53.5; DB 2; Length 46;
Best Local Similarity 48.0%; Pred. No. 2.3e+03;
Matches 12; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

Qy 209 PMPLPRQNHPEVA---TSPSSPPA 230
   :  :  |||  :  |||  :  |||  :  |||
Db 16 PAPVPEASTAPVAAPTTXPSPPAPA 40

RESULT 5
D81044
hypothetical protein NM1776 [imported] - Neisseria meningitidis (strain MC58 serogroup
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: D81044
F:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
  Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
  ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizzza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

```

A;Cross-references: UNIPROT:Q8XA85; GB:AE005174; NID:gl2512676; PIDN:AAG54301.1; GSPDB:G1
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: thrL

Query Match 2.8%; Score 48.5; DB 2; Length 27;
Best Local Similarity 44.4%; Pred. No. 2.8e+03;
Matches 12; Conservative 4; Mismatches 4; Indels 7; Gaps 1;

Qy 145 RTSTVTPTTTTPTTTVTPTTMSIPTTT 171
| |||: |||: |||: |||
Db 3 RISTTI-----TTTITTTITITTT 22

RESULT 11
AE3429
hypothetical protein BMEI1419 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AE3429
R;DeVeechio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Muler, C.; Los, T.; Ivanova, I.;
Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11755688
A;Accession: AE3429
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-47 <KUR>
A;Cross-references: UNIPROT:Q8YFV1; GB:AE008917; PIDN:AAL52600.1; PID:gl7983419; GSPDB:G1
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1419
A;Map position: 1

Query Match 2.7%; Score 47.5; DB 2; Length 47;
Best Local Similarity 45.9%; Pred. No. 5.6e+03;
Matches 17; Conservative 3; Mismatches 12; Indels 5; Gaps 3;

Qy 213 PRNHPEVATSP--SSQPASTHP-TTLOGAIRREPT 246
|||: | ||| |||: |||: |||
Db 12 PRR--DAVLRSPDPSAQPAATAPNTTTQNRWRELK 46

RESULT 12
I38970
fibroblast growth factor receptor 3 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004
C;Accession: I38970; S52862
R;Murgue, B.; Tsunekawa, S.; Rosenberg, I.; deBeaumont, M.; Podolsky, D.K.
Cancer Res. 54, 5206-5211, 1994
A;Title: Identification of a novel variant form of fibroblast growth factor receptor 3 (F)
A;Reference number: I38970; MUID:95007529; PMID:7923141
A;Accession: I38970
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-50 <RES>
A;Cross-references: UNIPROT:Q8NI16; EMBL:U22410; NID:9841313; PIDN:AAA67781.1; PID:9841313
R;Scotet, E.
A;Reference number: S52862
A;Accession: S52862
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-50 <SCO>
A;Cross-references: EMBL:X84939; NID:9695548; PID:9695549
C;Genetics:
A;Gene: GDB:FGFR3
A;Cross-references: GDB:127526; OMIM:100800; OMIM:134934
A;Map position: 4p16.3-4p16.3
A;Keywords: growth factor receptor

Query Match 2.7%; Score 47; DB 2; Length 50;
Best Local Similarity 45.0%; Pred. No. 6.3e+03;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Qy 87 DVSLTIENTAVSDSGVCCR 106
||| : : : : :
Db 10 DVRLRLANVSERDGGEXLCR 29

RESULT 13
F81142
hypotheical protein NMB0916 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: F81142
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: F81142
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <TET>
A;Cross-references: UNIPROT:Q9JZS5; GB:AE002443; GB:AE002098; NID:g7226149; PIDN:AAF4132
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0916

Query Match 2.6%; Score 46; DB 2; Length 46;
Best Local Similarity 30.0%; Pred. No. 6.7e+03;
Matches 12; Conservative 6; Mismatches 22; Indels 0; Gaps 0;

Qy 282 LLTANTTKGIYAGVCISVLVLLALIGVIAKKYFFKKEVQ 321
||| : : : : :
Db 4 LMAGVKRQIKQGGWLSVIALTSLFVSVFTLFYIFRHSVQ 43

RESULT 14
A85902
hypotheical protein Z3843 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: A85902
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A85902
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <STO>
A;Cross-references: UNIPROT:O8X447; GB:AE005174; NID:gi2516979; PIDN:AAG57677.1; GSPDB:C
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z3843

Query Match 2.6%; Score 45.5; DB 2; Length 50;
Best Local Similarity 27.7%; Pred. No. 7.8e+03;
Matches 13; Conservative 5; Mismatches 18; Indels 11; Gaps 1;

Qy 35 PCHYSGAVTSMCWNRGCSLFTCQNGIVMTNGTHVYRKDTRYKLLG 81
||| : : : : :
Db 6 PIETSGVVKALVWSAALDS-----NNSHVQFWGDVVRANLNG 41

RESULT 15
S20766
Ig heavy chain V region (VH4, IN109P1) - human (fragment)
C;Species: Homo sapiens (man)

C;Date: 20-Feb-1995 #sequence_revision 19-Apr-1996 #text_change 23-Jul-1999
C;Accession: S20766
R;Mortari, F.; Wang, J.; Schroeder, H.W.
submitted to the EMBL Data Library, April 1992
A;Description: Analysis of human cord blood Ig heavy chain IgA and IgG repertoire.
A;Reference number: S20764
A;Accession: S20766
A;Molecule type: DNA
A;Residues: 1-35 <MOR>
A;Cross-references: EMBL:Z11954; NID:g33861; PIDN:CAA78011.1; PID:g33862
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match 2.6%; Score 45; DB 2; Length 35;
Best Local Similarity 35.1%; Pred. No. 5.9e+03;
Matches 13; Conservative 6; Mismatches 10; Indels 8; Gaps 2;

Qy 89 SLTIENTAVSDSGVCCRVEH----RGWFNDMKITVS 121
||| : : : : :
Db 2 SVTVADTAV-----YYCARGKYGEGRRWGQGLTVTS 34

Search completed: June 29, 2005, 09:47:29
Job time : 44.6208 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:09:20 ; Search time 201.45 Seconds
(without alignments)
849.017 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLHLADSVAG.....PPKKEVQQLRPHKSCIHQRE 334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 68540

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60.5	3.5	49	2 Q8J4I5	Q8J4I5 chimpanzee
2	60	3.4	47	2 Q8J4H6	Q8J4H6 chimpanzee
3	59.5	3.4	44	2 Q8J4J3	Q8J4J3 chimpanzee
4	57	3.3	48	2 Q8J193	Q8J193 chimpanzee
5	57	3.3	48	2 Q8J256	Q8J256 chimpanzee
6	56.5	3.2	39	2 Q76N72	Q76N72 mus musculus
7	56	3.2	43	1 Y5K_C1MV	P16485 clover yell
8	55.5	3.2	50	2 Q8J273	Q8J273 chimpanzee
9	55	3.1	40	1 YC67_ARCFU	O29001 archaeoglob
10	55	3.1	47	2 Q8J174	Q8J174 chimpanzee
11	55	3.1	47	2 Q8J261	Q8J261 chimpanzee
12	54	3.1	42	2 Q8E825	Q8E825 shewanella
13	54	3.1	47	2 Q8J260	Q8J260 chimpanzee
14	54	3.1	49	2 Q8BBA2	Q8BBA2 human respi
15	53.5	3.1	46	2 Q7M1G3	Q7M1G3 lolium mult
16	53.5	3.1	46	2 Q8J4I7	Q8J4I7 chimpanzee
17	53	3.0	32	2 Q76N71	Q76N71 mus musculu
18	53	3.0	45	2 Q8J315	Q8J315 chimpanzee
19	53	3.0	48	2 Q8J164	Q8J164 chimpanzee
20	53	3.0	48	2 Q8J176	Q8J176 chimpanzee
21	53	3.0	48	2 Q8J194	Q8J194 chimpanzee
22	53	3.0	48	2 Q8J201	Q8J201 chimpanzee
23	53	3.0	48	2 Q8J222	Q8J222 chimpanzee
24	53	3.0	48	2 Q8J255	Q8J255 chimpanzee
25	53	3.0	48	2 Q8J257	Q8J257 chimpanzee
26	53	3.0	48	2 Q8J258	Q8J258 chimpanzee
27	53	3.0	48	2 Q8J265	Q8J265 chimpanzee
28	53	3.0	48	2 Q8J266	Q8J266 chimpanzee
29	53	3.0	48	2 Q8J269	Q8J269 chimpanzee
30	53	3.0	49	2 Q9PXH8	Q9PXH8 chimpanzee
31	52.5	3.0	47	2 Q6QIP3	Q6QIP3 hylobates m

ALIGNMENTS

RESULT 1

Q8J4I5 ID Q8J4I5 PRELIMINARY; PRT; 49 AA.
AC Q8J4I5;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Envelope glycoprotein (Fragment).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA Trichel A.M., Rajakumar P.A., Murphey-Corb M.A.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118212; AAM78572.1; -
DR GO; GO:0019031; C:Viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 49
SQ SEQUENCE 49 AA; 5101 MW; 86D35EEB15D4FC94 CRC64;

Query Match 3.5%; Score 60.5; DB 2; Length 49;
Best Local Similarity 50.0%; Pred. No. 4.5e+03;
Matches 17; Conservative 4; Mismatches 10; Indels 3; Gaps 2;

QY 164 TMSIPTTT-TVPTTMTVSTTTSVPTTTSIPTTTS 196
DB 4 TGNVPTTATTKTITTTTSTTT--PKADVNETSS 35

RESULT 2

Q8J4H6 ID Q8J4H6 PRELIMINARY; PRT; 47 AA.
AC Q8J4H6;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Envelope glycoprotein (Fragment).
OS Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11723;
RN [1]
RP SEQUENCE FROM N.A.
RA Trichel A.M., Rajakumar P.A., Murphey-Corb M.A.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118221; AAM78581.1; -
DR GO; GO:0019031; C:Viral envelope; IEA.
KW Envelope protein.
FT NON_TER 1
FT NON_TER 47
SQ SEQUENCE 47 AA; 4845 MW; EF30BF87DEF8C10B CRC64;

Query Match 3.4%; Score 60; DB 2; Length 47;


```

QY 176 TMTVSTTTSPVTTTSTPTTTSVPTVTTTSTFTV 207
Db 4 TGTAKTSTASPTTTA-----SSPSTTTVTPTSV 30

RESULT 9
YC67_ARCFU STANDARD; PRT; 40 AA.
ID YC67_ARCFU
AC Q29001;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein AF1267 precursor.
GN OrderedLocusNames=AF1267;
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RC MEDLINE=98049343; PubMed=9389475; DOI=10.1038/37052;
RA Klenk H.-P., Clayton R.A., Tomb J.-P., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M.L., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.R.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.J.,
RA Peterson S.N., Reich C.I., McNeil L.K., Badger J.H., Glodek A.,
RA Zhou L., Overbeek R., Gocayne J.D., Weidman J.F., McDonald L.A.,
RA Utterback T.R., Cotton M.D., Spriggs T., Artach P., Kaine B.P.,
RA Sykes S.M., Sadow P.W., D'Andrea K.P., Bowman C., Fujii C.,
RA Garland S.A., Mason T.M., Olsen G.J., Fraser C.M., Smith H.O.,
RA Woese C.R., Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE001017; AB89992.1; -.
DR F1R; B69408; B69408.
DR TIGR; AF1267; -.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1 17 Potential.
FT CHAIN 18 40 Hypothetical protein AF1267.
FT SEQUENCE 40 AA; 4306 MW; D2535A34416EC4E6 CRC64;
SQ
Query Match 3.1%; Score 55; DB 1; Length 40;
Best Local Similarity 28.6%; Pred. No. 8.3e+03;
Matches 14; Conservative 3; Mismatches 16; Indels 16; Gaps 2;

QY 17 SVAGSVKVGGEAGPSVTLCHYSGAVTSCWNRGSCSLTTCQGIWYTN 65
Db 2 AVAALAMYGTGCACAVLACN-----WNVREC-----GIITKN 34

RESULT 10
Q87174 PRELIMINARY; PRT; 47 AA.
ID Q87174
AC Q87174;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN Name=env;
GN Chimpantez immunodeficiency virus (SIV(cpz)) (CIV).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```



```
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Glycoprotein (Fragment).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22213634; PubMed=12226836; DOI=10.1002/jmv.10225;
RA Venter M., Collinson M., Schoub B.D.;
RT "Molecular epidemiological analysis of community circulating
RT respiratory syncytial virus in rural South Africa: Comparison of
RT viruses and genotypes responsible for different disease
RT manifestations.";
RL J. Med. Virol. 68:452-461(2002).
DR EMBL; AF548809; AAN62466.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0013062; P:virion attachment; IEA.
DR InterPro; IPR000925; Glycoprot_G.
DR Pfam; PF00802; Glycoprotein_G; 1.
FT NON_TER 1
FT SEQUENCE 49 AA; 5343 MW; DAB9CB5C89D596F CRC64;
SQ
Query Match 3.1%; Score 54; DB 2; Length 49;
Best Local Similarity 34.1%; Pred. No. 1.2e+04;
Matches 15; Conservative 4; Mismatches 25; Indels 0; Gaps 0;

QY 145 RSTSTVPTTTTPTTPTTMSIPTTTTPTTPTTMTVTTTTSVPTT 188
Db | | | | | | | | | | | | | | | | | | | | | |
2 RDPKTLAKTKKETTNPTKPTKPTKTRDSTSTQSVLDTATTS 45

RESULT 15
Q7M1G3 PRELIMINARY; PRT; 46 AA.
AC Q7M1G3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Arabinogalactan protein (Fragments).
OS Lolium multiflorum (Italian ryegrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Poaceae; Lolium.
OX NCBI_TaxID=4521;
RN [1]
RP SEQUENCE.
RX MEDLINE=90147544; PubMed=2695069;
RA Gleeson P.A., McNamara M., Wattenhall R.E.H., Stone B.A.,
RA Fincher G.B.;
RT "Characterization of the hydroxyproline-rich protein core of an
RT arabinogalactan-protein secreted from suspension-cultured Lolium
RT multiflorum (Italian ryegrass) endosperm cells.";
RL Biochem. J. 264:857-862(1989).
DR PIR; S07073; S07073.
DR NON_TER 1
DR NON_TER 46
FT NON_TER 46
FT SEQUENCE 46 AA; 4173 MW; EP61188B0C8A08D0 CRC64;
SQ
Query Match 3.1%; Score 53.5; DB 2; Length 46;
Best Local Similarity 48.0%; Pred. No. 1.2e+04;
Matches 12; Conservative 2; Mismatches 8; Indels 3; Gaps 1;

QY 209 FNPPLRQNHPEVA---TSPSPQPA 230
Db | | | | | | | | | | | | | | | | | | | | | |
16 PAPVPEASTAPVAAPTTPSPAPA 40
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:27:45 ; Search time 6.94695 Seconds
(without alignments)
720.211 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300
Perfect score: 275
Sequence: 1 LYSYTTDNDVTVESSDGL.....SLLTANTTKGIYAGVCIISVL 52
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 11837
Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43.5	15.8	49	2 S49365	Ig mu chain transmembrane region - bowfin
2	42.5	15.5	46	2 A49281	pol protein - simi
3	42	15.3	50	2 B49410	t-complex polypept
4	39	14.2	30	2 D82251	hypothetical prote
5	38	13.8	24	2 S00881	T-cell receptor al
6	38	13.8	35	2 S30112	aminoglycoside Ng'
7	37	13.5	34	2 S68880	ribulose-bisphosph
8	37	13.5	41	2 S38918	hypothetical prote
9	37	13.5	47	2 AF0464	hypothetical prote
10	37	13.5	48	2 S29150	superoxide dismuta
11	36.5	13.3	46	2 A49281	pol protein - simi
12	36	13.1	39	2 AC0205	hypothetical prote
13	35.5	12.9	29	2 A32860	biotin-binding pro
14	35.5	12.9	33	2 I46601	myosin - pig (frag
15	35.5	12.9	33	2 I46600	myosin - pig (frag
16	35.5	12.9	37	2 AG0540	hypothetical prote
17	35	12.7	50	2 A61149	endo-1,4-beta-xyla
18	34.5	12.5	46	2 A49281	pol protein - simi
19	34.5	12.5	48	2 PL0090	Ig heavy chain v r
20	34	12.4	30	2 S70341	napin large chain
21	34	12.4	31	2 S49191	hypothetical prote
22	34	12.4	40	2 PQ0202	endo-1,4-beta-xyla
23	34	12.4	41	2 H82525	hypothetical prote
24	33.5	12.2	40	2 H95183	hypothetical prote
25	33	12.0	24	2 A36912	hypothetical prote
26	33	12.0	38	2 S72344	pili protein - Nei
27	33	12.0	41	2 C30208	hypothetical prote
28	33	12.0	48	2 AC1046	entericidin B prec
29	33	12.0	48	2 H66110	hypothetical prote

30	32.5	11.8	31	2 H82682	hypothetical prote
31	32.5	11.8	36	2 B42418	auracyanin A - Chl
32	32.5	11.8	46	2 E49281	pol protein - simi
33	32	11.6	26	2 B53113	Lys-gingipain form
34	32	11.6	28	2 G32351	34K class B flagel
35	32	11.6	29	2 I50214	protein-cytosine-p
36	32	11.6	31	2 A57001	endo-1,4-beta-xyla
37	32	11.6	41	2 S56120	collagenase type 1
38	32	11.6	42	2 I37543	MHC HLA-DR-beta-1
39	32	11.6	45	2 D45731	gene 39.2 protein
40	32	11.6	48	2 AC1889	hypothetical prote
41	31.5	11.5	46	2 T43389	hypothetical prote
42	31	11.3	16	2 F45039	T-cell receptor be
43	31	11.3	19	2 I49422	L-lactate dehydrog
44	31	11.3	23	2 S47188	T-cell receptor J-
45	31	11.3	26	2 F45087	cysteine proteinas

ALIGNMENTS

RESULT 1

S49365

Ig mu chain transmembrane region - bowfin

C:Species: Amia calva (bowfin)

C>Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 20-Mar-1998

C:Accession: S49365

R:Wilson, M.R.; Ross, D.A.; Miller, N.W.; Clem, L.W.; Middleton, D.L.; Warr, G.W.

A:Description: Alternate pre-mRNA splicing pathways in the production of membrane IgM he

A:Reference number: S49365

A:Accession: S49365

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-49 <WIL>

A:Cross-references: EMBL:X82070; NID:g557218; PID:g557219

Query Match 15.8%; Score 43.5; DB 2; Length 49;

Best Local Similarity 34.5%; Pred. No. 1.4e+02;

Matches 10; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

QY 2 LYSYTTDNDVTVESSDGLWNNQTLFL 30

DB 5 VFQYTVQDNG-IEYNGDSLWNTVCTFIFL 32

RESULT 2

A49281

pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (isolate Bab34) (fragment)

C:Species: simian T-cell lymphotropic virus type 1, STLV-1

C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004

C:Accession: A49281

R:Sakaena, N.K.; Herve, V.; Durand, J.P.; Leguenno, B.; Diop, O.M.; Digoutte, J.P.; Mathi

; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Poiesz, B.J.

Virology 198, 297-310, 1994

A:Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukemia

A:Reference number: A49281; MUID:94082462; PMID:8259665

A:Accession: A49281

A:Status: preliminary; translation not shown

A:Molecule type: DNA

A:Residues: 1-46 <SAK>

A:Cross-references: UNIPROT:Q88390; GB:L20351; NID:g431689; PIDN:AAA16596.1; PID:g431690

C:Superfamily: pol polyprotein

Query Match

Best Local Similarity 15.5%; Score 42.5; DB 2; Length 46;

Matches 13; Conservative 3; Mismatches 14; Indels 3; Gaps 2;

QY 2 LYSYTTDNDVTVESSDGLWNNQTLFL 32

DB 13 LYKFTDPLDLPMDNALSIGLWTFNHLNV-LTH 44


```
RESULT 8
S38918
hypothetical protein 6 - phage phi-C31
C:Species: phage phi-C31
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004
R:Accession: S38918
R:Hartley, N.M.; Murphy, G.O.; Bruton, C.J.; Chater, K.P.
submitted to the EMBL Data Library, November 1993
A:Reference number: S38912
A:Accession: S38918
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-41 <HAR>
A:Cross-references: UNIPROT:Q38027; EMBL:X76288; NID:G432610; PIDN:CAA53917.1; PID:G5790
C:Genetics:
A:Start codon: GTG

Query Match 13.5%; Score 37; DB 2; Length 41;
Best Local Similarity 45.5%; Pred. No. 7.1e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 42 GYAGVCISVL 52
| : | : | : |
DB 12 GWMVGICVGL 22

RESULT 9
AF0464
hypothetical protein YPO3815 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AF0464
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
N., M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AF0464
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-47 <KUR>
A:Cross-references: UNIPROT:Q82A17; GB:AL590842; PIDN:CAC93282.1; PID:G15981729; GSPDB:G
C:Genetics:
A:Gene: YPO3815

Query Match 13.5%; Score 37; DB 2; Length 47;
Best Local Similarity 39.1%; Pred. No. 8.3e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 2; Gaps 1;

QY 17 SDGLWNNQTLFLEHSLTANT 39
::|||::|::|::|::|
DB 8 NNSLGYNNVQVY--HFLDSNT 28

RESULT 10
S29150
superoxide dismutase (EC 1.15.1.1) (Cu-Zn) III, cytosolic - rice (fragment)
C:Species: Oryza sativa (rice)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: S29150; B61531
R:Kanematsu, S.; Aaada, K.
Plant Cell Physiol 31, 99-112, 1990
A:Title: Characteristic amino acid sequences of chloroplast and cytosol isozymes of CuZn
A:Reference number: S29146
A:Accession: S29150
A:Molecule type: protein
A:Residues: 1-48 <KAN>
A:Cross-references: UNIPROT:Q7M237
R:Kanematsu, S.; Aaada, K.
Free Radic. Res. Commun. 12, 383-390, 1991
A:Title: Chloroplast and cytosol isozymes of CuZn-superoxide dismutase: their characteri
```

```
A:Reference number: A61531
A:Accession: B61531
A:Molecule type: protein
A:Residues: 1-48 <KA2>
C:Function:
A:Description: catalyzes the dismutation of 2 molecules of peroxide radical to dioxygen
C:Superfamily: superoxide dismutase (Cu-Zn)
C:Keywords: copper; cytosol; metalloprotein; oxidoreductase; zinc
F:44.46/Binding site: copper (His) #status predicted

Query Match 13.5%; Score 37; DB 2; Length 48;
Best Local Similarity 50.0%; Pred. No. 8.5e+02;
Matches 9; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

QY 3 YSYTTDGNQDVTVTSSDGL 20
:| ||| ||| |||
DB 19 FSQEGDGPTTXXSVSGL 36

RESULT 11
F49281
pol protein - simian T-cell lymphotropic virus type 1, STLV-1 (fragment)
C:Species: simian T-cell lymphotropic virus type 1, STLV-1
C>Date: 19-Mar-1997 #sequence_revision 19-Dec-1997 #text_change 09-Jul-2004
C:Accession: F49281; B49281
R:Saksena, N.K.; Herve, V.; Durand, J.P.; Leguennou, B.; Diop, O.M.; Digoutte, J.P.; Mathi
; Paul, B.; Dube, D.K.; Barre-Sinoussi, F.; Polesz, B.J.
Virology 198, 297-310, 1994
A:Title: Seroepidemiologic, molecular, and phylogenetic analyses of simian T-cell leukemia
A:Reference number: A49281; MUID:94082462; PMID:8259665
A:Accession: F49281
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-46 <SAK>
A:Cross-references: UNIPROT:Q88344; UNIPROT:Q88360; UNIPROT:Q88343; UNIPROT:Q88394; UNIP
PROT:Q88331; UNIPROT:Q88406; GB:I20357; NID:G431721; PIDN:AAA16561.1; PID:G431722
A:Experimental source: isolate SAB37Pol
A:Accession: B49281
A>Status: preliminary; translation not shown
A:Molecule type: DNA
A:Residues: 1-46 <SA2>
A:Cross-references: GB:L20358; NID:G431709; PIDN:AAA16556.1; PID:G431710
A:Experimental source: isolate SAB91037
C:Superfamily: pol polyprotein

Query Match 13.3%; Score 36.5; DB 2; Length 46;
Best Local Similarity 36.4%; Pred. No. 9.3e+02;
Matches 12; Conservative 3; Mismatches 15; Indels 3; Gaps 2;

QY 2 LYSYTTDGNQDVTVTSS--SDGLWNNQTLFLEH 32
||| ||| ||| :| ||| |||
DB 13 LYKYFTDKPDLPMDNALSIALWTINHLNV-LTH 44

RESULT 12
AC0205
hypothetical protein YPO1682 [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AC0205
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AC0205
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-39 <KUR>
A:Cross-references: UNIPROT:Q8ZFL7; GB:AL590842; PIDN:CAC90503.1; PID:G15979714; GSPDB:G
C:Genetics:
A:Gene: YPO1682
```


Result No.	Query	Score	Query			DB	ID	Description
			Match	Length	Length			
1	47	17.1	43	2	Q86LU3	cryptospori		
2	46	16.7	43	2	Q81OR6	cryptospori		
3	46	16.7	43	2	Q81OR7	cryptospori		
4	46	16.7	43	2	Q81SB4	cryptospori		
5	45.5	16.5	48	2	Q6WV04	saccharomyc		
6	44	16.0	43	2	Q86LU1	cryptospori		
7	44	16.0	43	2	Q86LU6	cryptospori		
8	44	16.0	46	2	Q8J4I8	chimpanzee		
9	44	16.0	49	2	Q8J4I5	chimpanzee		
10	43.5	15.8	41	2	Q99J74	mus musculus		
11	43	15.6	41	2	Q7RDL1	plasmodium		
12	43	15.6	43	2	Q86LU2	cryptospori		
13	43	15.6	43	2	Q86LU5	cryptospori		
14	42.5	15.5	46	2	Q88390	simian t-ly		
15	42	15.3	43	2	Q86LU4	cryptospori		
16	42	15.3	46	2	Q7KYS6	homo sapien		
17	42	15.3	47	2	Q8J4I9	chimpanzee		
18	41.5	15.1	47	2	Q75JX4	dictyosteli		
19	41	14.9	29	2	Q86233	human rotav		
20	41	14.9	40	2	Q83DR8	coxiella bu		
21	41	14.9	41	2	O11552	human immun		
22	40.5	14.7	49	2	Q91FZ1	chilo iride		
23	40	14.5	49	2	Q7XA54	gnemum gnem		
24	39.5	14.4	46	2	Q82209	human t-lym		
25	39.5	14.4	47	2	Q89Z69	leptospira		
26	39	14.2	28	2	Q998S5	turkey ente		
27	39	14.2	30	2	Q9K7V5	vibrio chol		
28	39	14.2	35	2	Q6QPN9	human immun		
29	39	14.2	38	2	Q7RD73	plasmodium		
30	39	14.2	47	2	Q65SH0	mannheimia		
31	39	14.2	50	2	Q7SEZ1	neurospora		

RN	[1]
RP	SEQUENCE FROM N.A.
RA	Trichel A.M., Rajakumar P.A., Murphey-Corb M.A.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY118209; AAU78569.1; -
DR	GO; GO:0019031; C:viral envelope; IEA.
KW	Envelope protein.
FT	NON_TER 1
FT	NON_TER 46
FT	NON_TER 46
SQ	SEQUENCE 46 AA; 4928 MW; 9A2C139D4777658E CRC64;
 Query Match 16.0%; Score 44; DB 2; Length 46; Best Local Similarity 48.0%; Pred. No. 6.2e+02; Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;	
QY	4 SYTTDGNNTVTSSDGLMNNOTQL 28
DG	18 STTTPKADVVNETSSCVKNNTCTGL 42
 NCBI_TaxID=11723;	
RESULT 9	
Q8J4I5	PRELIMINARY; PRT; 49 AA.
ID	Q8J4I5
AC	Q8J4I5;
DT	01-OCT-2002 (TEMBLrel. 22, Created)
DT	01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE	Envelope glycoprotein (Fragment).
OS	Chimpanzee immunodeficiency virus (SIV(cpz)) (CIV).
OC	Viruses; Retrod viruses; Retroviridae; Lentivirus.
OX	NCBI_TaxID=11723;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Trichel A.M., Rajakumar P.A., Murphey-Corb M.A.;
RL	Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY118212; AAU78572.1; -
DR	GO; GO:0019031; C:viral envelope; IEA.
KW	Envelope protein.
FT	NON_TER 1
FT	NON_TER 49
FT	NON_TER 49
SQ	SEQUENCE 49 AA; 5101 MW; 86D3EEB15D4FC94 CRC64;
 Query Match 16.0%; Score 44; DB 2; Length 49; Best Local Similarity 48.0%; Pred. No. 6.7e+02; Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;	
QY	4 SYTTDGNNTVTSSDGLMNNOTQL 28
DG	21 STTTPKADVVNETSSCVKNNTCTGL 45
 NCBI_TaxID=10090;	
RESULT 10	
Q99JT4	PRELIMINARY; PRT; 41 AA.
ID	Q99JT4
AC	Q99JT4;
DT	01-JUN-2001 (TEMBLrel. 17, Created)
DT	01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT	01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE	Hypothetical protein.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RR	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA	Aklauser R.L., Feingold E.A., Grouse L.H., Derge J.G.
RA	Klausner R.D., Collins F.S., Wagner L., Shennen C.W., Schuler G.D.,
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA	Hopkins R.F., Jordan H., Moore T.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Scapleton M., Soares M.B., Ronaldo M.F., Casavant T.L., Scheetz T.E.

RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA	Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA	Jones S.J., Marra M.A.;
RT	"Generation and initial analysis of more than 15,000 full-length human
RT	and mouse cDNA sequences.";
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA	Strausberg R.;
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC005705; AAH05705.1; -.
KW	Hypothetical protein.
SQ	SEQUENCE 41 AA; 4720 MW; 7692D458655D3B05 CRC64;
Query Match	15.8%; Score 43.5; DB 2; Length 41;
Best Local Similarity	31.6%; Pred. No. 6.3e+02;
Matches 12; Conservative 3; Mismatches 18; Indels 5; Gaps 1;	
Qy	2 LYSYTTDGNITVTESSDGLNWNQTQLFLEHSLLTANT 39
DB	3 LYLLRLDANHTIL-----LMSKNYKCLGTNHGLKVARS 35
RESULT 11	
Q7RD11	PRELIMINARY; PRT; 41 AA.
ID Q7RD11	
AC Q7RD11	
DT 01-MAR-2004 (TREMBLrel. 26, Created)	
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)	
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)	
DE Hypothetical protein (Fragment).	
GN Name=PY05411;	
OS Plasmodium yoelii yoelii.	
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.	
OX NCBI_TaxID=73239;	
PI [1]	
RN SEQUENCE FROM N.A.	
RC STRAIN=17XNL;	
PUBMed=12368865; DOI=10.1038/nature01099;	
RA Carlton J.M., Angiuoli S.V., Suh B.B., Koelij T.W., Pertea M.,	
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,	
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,	
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblum T.V.,	
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,	
RA Florens L., Yates F.R. III, Raine J.D., Sindén R.E., Harris M.A.,	
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,	
van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,	
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,	
RA Carucci D.J.;	
RT "Genome sequence and comparative analysis of the model rodent malaria	
RT parasite Plasmodium yoelii yoelii.";	
RL Nature 419:512-519(2002).	
CC -! CAUTION: The sequence shown here is derived from an	
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is	
CC preliminary data.	
DR EMBL; AABL01001711; EAAI7433.1; -.	
KW Hypothetical protein.	
FT NON TER 1	
SQ SEQUENCE 41 AA; 4786 MW; E19A7D61ABFB24F0 CRC64;	
Query Match	15.6%; Score 43; DB 2; Length 41;
Best Local Similarity	42.9%; Pred. No. 7.3e+02;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;	

```

Db      : || : : : || | : | : ||
3 FVTDLSSDLSNTADGLQAW-YKISQAEIEHIMTA 37

```

Search completed: June 29, 2005, 09:46:16
Job time : 32.3635 secs

Db 3 FVTDLSSDLNTADGLQAW-YKISQVEVEHIIMTA 37

[illegible]

Query Match 15.5%; Score 42.5; DB 2; Length 46;
Best Local Similarity 39.4%; Pred. No. 9.5e+02;
Matches 13; Conservative 3; Mismatches 14; Indels 3; Gaps 2;

QY 2 LYSYTTDGNIVTES--SDGLWNNQTQLFLEH 32
||| ||| : ||| : |||
Db 13 LYKYFTDKPILPMDNALSIGLWTFEHLNV-LTH 44

RESULT	15
Q86LU4	
ID	Q86LU4 PRELIMINARY; PRT; 43 AA.
AC	Q86LU4;
DT	01-JUN-2003 (TrEMBLrel. 24, Created)
DT	01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE	01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE	Hypochemical protein (Fragment).
OS	Cryptosporidium felis.
OC	Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC	Cryptosporidiidae; Cryptosporidium.
OX	NCBI_TaxID=83540;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=1626;
RC	MEDLINE=22511025; PubMed=12624019;
RA	Leoni F., Gallimore C.I., Green J., McLauchlin J.;
RT	"Molecular epidemiological analysis of Cryptosporidium isolates from humans and animals by using a heteroduplex mobility assay and nucleic acid sequencing based on a small double-stranded RNA element.";
RL	J. Clin. Microbiol. 41:981-992(2003).
DR	EMBL; AV176651; AAO19456.1; -.
KW	Hypochemical protein.
FT	NON_TER 1
FT	NON_TER 43
SO	SEQUENCE 43 AA; 4798 MW; CB1F81C40B323833 CRC64;

```

Query Match      15.3%; Score 42; DB 2; Length 43;
Best Local Similarity 30.6%; Pred. No. 1e+03;
Matches 11: Conservative 10; Mismatches 11; Indels 4; Gaps 2;

```

QY 5 YTTDGN DTVT ESDGL--WNNNQTLFLEHSLTA 37

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 36.9823 Seconds
(without alignments)

543.815 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300

Perfect score: 275

Sequence: 1 PLYSVTTDGNDRVTSSDGL.....SLLTANTTKGIYAGVCISVL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003Bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	275	100.0	334	2	Aaw38336 Human kid
2	275	100.0	334	6	Aao26679 KIM-1 rel
3	275	100.0	334	8	Adq76696 Human kid
4	275	100.0	339	7	Adq76696 Plasmid p
5	275	100.0	339	7	Adq76696 Human NOV
6	275	100.0	359	6	Adr58582 Human can
7	275	100.0	359	6	Abp70439 Amino aci
8	275	100.0	359	6	Abp70438 Amino aci
9	275	100.0	359	6	Abp70438 Human bla
10	275	100.0	359	6	Aao26680 KIM-1 rel
11	275	100.0	359	7	Adq76696 Human NOV
12	275	100.0	359	7	Adn38984 Cancer/an
13	275	100.0	359	8	Adq76696 Human kid
14	275	100.0	364	6	Abp70442 Amino aci
15	275	100.0	365	6	Abp70440 Amino aci
16	270	98.2	359	6	Abp70441 Amino aci
17	270	98.2	364	6	Abp70443 Amino aci
18	228	82.9	298	8	Adq76693 Human KIM
19	228	82.9	451	2	AAR92803 Hepatitis
20	228	82.9	518	8	Adq76691 Human KIM
21	225	81.8	81	6	Aao26678 Human KIM
22	189	68.7	263	7	Adq76692 Human NOV
23	189	68.7	263	7	Adq76692 Plasmid p
24	189	68.7	263	7	Adq76696 Human NOV
25	101	36.7	18	6	Aao26686 Human KIM

26	99	36.0	18	6	Aao26687	Aao26687 Human KIM
27	99	36.0	18	6	Aao26673	Aao26673 Monoclonal
28	90	32.7	18	6	Aao26688	Aao26688 Human KIM
29	82	29.8	18	6	Aao26685	Aao26685 Human KIM
30	79	28.7	305	6	Abp70430	Abp70430 Amino aci
31	78	28.4	307	2	Aaw38334	Aaw38334 Rat kidne
32	76.5	27.8	282	6	Abp70431	Abp70431 Amino aci
33	62.5	22.7	774	8	Adg93320	Adg93320 DEN1 (Pue
34	62.5	22.7	775	8	Adg93318	Adg93318 DEN1 (Pue
35	62	22.5	606	4	Abp78898	Abp78898 C. glutam
36	60.5	22.0	474	5	Abp04708	Abp04708 Human SP1
37	60.5	22.0	474	6	Abp00247	Abp00247 Human nov
38	60.5	22.0	474	7	Adg31674	Adg31674 Human pro
39	60	21.8	219	3	Aag19395	Aag19395 Arabidops
40	60	21.8	249	3	Aag53467	Aag53467 Arabidops
41	60	21.8	249	3	Aag19394	Aag19394 Arabidops
42	60	21.8	268	3	Aag53466	Aag53466 Arabidops
43	60	21.8	268	3	Aag19393	Aag19393 Arabidops
44	60	21.8	338	3	Aag53465	Aag53465 Arabidops
45	60	21.8	338	8	Ado02101	Ado02101 Thalecres

ALIGNMENTS

RESULT 1

AAW38336

ID AAW38336 standard; protein; 334 AA.

XX AAW38336;

XX AAW38336;

DT 21-MAY-1998 (first entry)

DE Human kidney injury related molecule (KIM).

KW Kidney injury related molecule; KIM; human; renal disease; injury; nephritis; tissue regeneration; therapy; monoclonal antibody.

OS Homo sapiens.

PN WO744460-A1.

PD 27-NOV-1997.

PF 23-MAY-1997; 97WO-US009303.

PR 24-MAY-1996; 96US-0018228P.

PR 23-AUG-1996; 96US-0023442P.

XX (BIOJ) BIOGEN INC.

PI Sanicola-Nadel M, Bonventre JV, Hession CA, Ichimura T, Wei H; Cate RL;

XX WPI; 1998-018514/02.

DR DNA encoding kidney injury related molecule - which is upregulated in injured or regenerating tissue, useful to promote growth of new tissue and survival of damaged tissue.

PT Claim 9; Page 46-47; 68pp; English.

PS This protein, designated kidney injury related molecule (KIM), is up-regulated in injured or regenerating tissue. Its amino acid sequence was deduced from a clone (see AAT96035) obtained from a human embryonic liver library. A 577-amino acid (see AAW38335) and a 307-amino acid rat KIM (see AAW38334) are also claimed. Recombinant KIM polypeptides can be expressed in prokaryotic and eukaryotic host cells using a claimed process. Soluble variants fused to a toxin, imageable compound or radionuclide, and IgG fusion proteins are also claimed. KIM, or an agonist, can be used to treat renal disease and to promote the growth of new tissue or the survival of damaged tissue, generally in conditions where the binding of specific ligand to KIM stimulates cell growth, maintains

CC cellular differentiation or reduces apoptosis, e.g. in cases of renal
CC failure, nephritis, kidney transplants, toxic or hypoxic injury. A
CC monoclonal antibody specific for KIM can be used to treat renal disease,
CC e.g. where binding of KIM to ligand results in neoplasia, loss of
CC cellular function, susceptibility to apoptosis or promotion of
CC inflammation, deliver imaging agents to KIM expressing cells in vivo or
CC in vitro and measure KIM concentration by immunoassay.
CC Damage/regeneration of renal cells can be determined by measuring KIM,
CC particularly to diagnose or monitor the progress of disease or therapy.
CC KIM-expressing tumour cells can be inhibited by treatment with a fusion
CC protein comprising KIM ligand or MAb with a toxin or radionuclide, and
CC tumour cells that express KIM ligand can be inhibited with similarly
CC tagged KIM or anti-KIM ligand antibody
XX
SQ Sequence 334 AA;

Query Match 100.0%; Score 275; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 7.3e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSVTTDGNVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
|||||
DB 249 PLYSVTTDGNVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 2
AA026679
ID AAO26679 standard; protein; 334 AA.

AC AAO26679;
DT 20-MAR-2003 (first entry)
XX
DE KIM-1 related protein, SEQ ID No 7.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX
OS Homo sapiens.

XX WO200298920-A1.

PN 12-DEC-2002.

PD 31-MAY-2002; 2002WO-US017402.

PF 01-JUN-2001; 2001US-0295449P.

PR 04-JUN-2001; 2001US-0295907P.

XX (BIOJ) BIOGEN INC.

PA (GEO) GEN HOSPITAL CORP.

XX Bailly V, Bonventre J;

XX WPI; 2003-156845/15.

XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.

XX Disclosure; Page 39-40; 42pp; English.

XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a protein relating to the KIM-1
CC protein of the invention

SQ Sequence 334 AA;

Query Match 100.0%; Score 275; DB 6; Length 334;
Best Local Similarity 100.0%; Pred. No. 7.3e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSVTTDGNVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
|||||
DB 249 PLYSVTTDGNVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 3
ADQ76696
ID ADQ76696 standard; protein; 334 AA.

XX ADQ76696;

DT 16-DEC-2004 (first entry)

XX Human kidney injury molecule-1 (KIM-1).

XX Human; KIM-1; kidney injury molecule-1; immunosuppressive; protozoacide;
KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW antiallergic; antiinflammatory; vasotropic; nephropic.

XX Homo sapiens.

FH Key Location/Qualifiers
FT Peptide 1..20
FT Modified-site /label= Signal_peptide
FT Modified-site /note= "N-glycosylated"

FT Modified-site 258..260
FT Modified-site /note= "N-glycosylated"

FT Modified-site 272..274
FT Modified-site /note= "N-glycosylated"

FT Modified-site 286..288
FT Modified-site /note= "N-glycosylated"

FT Domain 290..311
FT /note= "Transmembrane domain"

FT Region 324..334

FT /note= "C-terminal region unique to this splice variant"

XX WO2004060041-A2.

XX 22-JUL-2004.

XX 29-DEC-2003; 2003WO-US041294.

XX 30-DEC-2002; 2002US-0436934P.

XX (BIOG-) BIOGEN IDEC MA INC.

XX Rennert PD;

XX WPI; 2004-534277/51.

XX Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
PT Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
PT activation, or inhibiting production of a subset of antibodies against
PT one or more antigens.

XX Disclosure; SEQ ID NO 1; 68pp; English.

XX The present sequence is that of human kidney injury molecule-1 (KIM-1), a
CC type I cell membrane glycoprotein and member of the T cell immunoglobulin
CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids or
CC 359 amino acids ADQ76690, depending on splice variation. It has been
CC discovered that treatment of a mammal with a KIM-1 antagonist alters the
CC interaction of T cells and other immune system cells, e.g. dendritic
CC cells, monocytes, macrophages and B cells, and thereby strongly
CC suppresses an IgG response to an antigen. Such treatment also eliminates
CC IgG1 production by memory B cells in response to subsequent challenge

CC with the antigen. Blockage of the binding of KIM-1 to its receptor
 CC reduces secretion of interferon-gamma by immune cells engaged in an
 CC antigen response in the mixed lymphocyte response assay. Based on these
 CC discoveries, the invention provides methods for therapeutically
 CC modulating immune function in autoimmune diseases and other disorders of
 CC the mammalian immune system. A method is claimed for inhibiting
 CC signalling between a T cell and a second cell, e.g. an antigen-presenting
 CC cell, in a mammal. The method involves identifying a mammal with an
 CC immune disease or disorder, or one preparing to receive a tissue graft,
 CC and administering a KIM-1 antagonist, especially a polypeptide
 CC comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a
 CC KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding
 CC fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a
 CC soluble polypeptide, which can include a KIM-1 mucin domain in addition
 CC to the KIM Ig domain, and may include a heterologous moiety such as an Fc
 CC moiety. The antagonist may be conjugated to a polymer such as
 CC polyethylene glycol. The KIM-1 antagonist is used in methods of:
 CC inhibiting activation of a B cell in a mammal; inhibiting disease relapse
 CC in an autoimmune disease; inhibiting epitope spreading in an autoimmune
 CC disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis,
 CC autoimmune haemolytic anaemia, Chagas disease, Grave's disease,
 CC idiopathic thrombocytopenia purpura, Wegener's granulomatosis,
 CC polyarthritis nodosa, rapidly progressive crescentic glomerulonephritis,
 CC graft-versus-host disease, or systemic lupus nephritis) or inflammatory
 CC disease or disorder (e.g. inflammatory bowel diseases such as Crohn's
 CC disease, ulcerative colitis, and ileitis); and inhibiting secretion of
 CC IFN-gamma by lymphocytes.

XX
 SQ Sequence 334 AA;

Query Match 100.0%; Score 275; DB 8; Length 334;
 Best Local Similarity 100.0%; Pred. No. 7.3e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNNTQFLHSLLTANTTKGIYAGVCISVL 52
 DB 249 PLYSYTTDGNVTVTSSDGLWNNNTQFLHSLLTANTTKGIYAGVCISVL 300

RESULT 4
 ADE36618
 ID ADE36618 standard; protein; 339 AA.

XX ADE36618;
 XX
 XX 29-JAN-2004 (first entry)
 XX
 DE Plasmid pCR2.1-CGS7008-03-S843 15B protein insert SEQ ID NO:28.
 XX
 KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
 KW renal cancer; inflammation; tissue typing.
 XX
 OS Synthetic.
 OS Homo sapiens.
 PN WO2003080856-A2.
 XX
 PD 02-OCT-2003.
 XX
 PF 19-MAR-2003; 2003WO-US008490.
 XX
 PR 19-MAR-2002; 2002US-0365491P.
 PR 13-SEP-2002; 2002US-0410618P.
 XX
 XX (CURA-) CURAGEN CORP.

XX Lepley DM, Rieger DK, Tse X, Rastelli L, Smithson G, Meseri M;
 PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;
 XX
 XX WPI: 2003-876927/81.
 DR N-PSDB; ADE36617.
 XX
 XX New polypeptide, useful for preparing a composition for treating or

PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
 PT or inflammation, or for tissue typing.
 XX
 PS Example 1; SEQ ID NO 28; 239pp; English.

CC The present invention describes an isolated human NOVX polypeptide, where
 CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
 CC least 95 % identical to it, or a sequence comprising one or more
 CC conservative substitutions in the amino acid sequence. The NOVX
 CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
 CC antiinflammatory activities, and can be used in gene therapy, and in
 CC vaccines. The NOVX polypeptide is useful for preparing a composition for
 CC treating or preventing a pathology associated with NOVX polypeptide e.g.
 CC renal cancer or inflammation, or for tissue typing. The present sequence
 CC represents a pCR2.1-CGS7008-03-S843 15B protein insert, which is used in
 CC an example from the present invention.

XX Sequence 339 AA;

Query Match 100.0%; Score 275; DB 7; Length 339;
 Best Local Similarity 100.0%; Pred. No. 7.5e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNNTQFLHSLLTANTTKGIYAGVCISVL 52
 DB 229 PLYSYTTDGNVTVTSSDGLWNNNTQFLHSLLTANTTKGIYAGVCISVL 280

RESULT 5
 ADE36592
 ID ADE36592 standard; protein; 339 AA.

XX ADE36592;
 XX
 XX 29-JAN-2004 (first entry)
 XX
 DE Human NOV1a protein SEQ ID NO:2.
 XX
 KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
 KW renal cancer; inflammation; tissue typing.

OS Homo sapiens.
 XX
 PN WO2003080856-A2.
 XX
 PD 02-OCT-2003.

XX
 PF 19-MAR-2003; 2003WO-US008490.
 XX
 PR 19-MAR-2002; 2002US-0365491P.
 PR 13-SEP-2002; 2002US-0410618P.
 XX

PA (CURA-) CURAGEN CORP.

XX Lepley DM, Rieger DK, Tse X, Rastelli L, Smithson G, Meseri M;
 PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;

XX WPI: 2003-876927/81.
 DR N-PSDB; ADE36591.

XX New polypeptide, useful for preparing a composition for treating or
 PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
 PT or inflammation, or for tissue typing.

PS Claim 1; SEQ ID NO 2; 239pp; English.

CC The present invention describes an isolated human NOVX polypeptide, where
 CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
 CC least 95 % identical to it, or a sequence comprising one or more
 CC conservative substitutions in the amino acid sequence. The NOVX
 CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
 CC antiinflammatory activities, and can be used in gene therapy, and in
 CC vaccines. The NOVX polypeptide is useful for preparing a composition for

CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOV1a from the present invention.

XX
SQ Sequence 339 AA;

```
Query Match      100.0%; Score 275; DB 7; Length 339;
Best Local Similarity 100.0%; Pred. No. 7.5e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	P L Y S Y T T D G N D V T V E S D G L W N N O T Q L F L E H S I L T A N T T K G I Y A G V C I S V L	52
D_b	229	P L Y S Y T T D G N D V T V E S D G L W N N O T Q L F L E H S I L T A N T T K G I Y A G V C I S V L	28

CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies

Sequence 359 AA:

Query Match	100.0%;	Score 275;	DB 6;	Length 359;
Best Local Similarity	100.0%;	Pred. No. 8e-26;		
Matches 52:	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;

QY	1	PLYSYTTDGN	DTVT	ESSDGL	WNNQ	TQ	LF	HS	LL	TANT	KG	IY	AG	VC	IS	VL	52
pb	249	PLYSYTTDGN	DTVT <td>ESSDGL <td>WNNQ <td>TQ <td>LF <td>HS <td>LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td></td></td></td></td></td></td>	ESSDGL <td>WNNQ <td>TQ <td>LF <td>HS <td>LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td></td></td></td></td></td>	WNNQ <td>TQ <td>LF <td>HS <td>LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td></td></td></td></td>	TQ <td>LF <td>HS <td>LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td></td></td></td>	LF <td>HS <td>LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td></td></td>	HS <td>LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td></td>	LL <td>TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td></td>	TANT <td>KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td></td>	KG <td>IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td></td>	IY <td>AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td></td>	AG <td>VC <td>IS <td>VL</td> <td>300</td> </td></td>	VC <td>IS <td>VL</td> <td>300</td> </td>	IS <td>VL</td> <td>300</td>	VL	300

RESULT 7
APPENDIX 39

ABP70439 standard; protein; 359 AA.

XX
AC ABP70439;

DT 22-APR-2003 (first entry)

XX Amino acid sequence of human TIM-1 allele 1.

XX
KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.

XX
OS Homo sapiens.XX
PN WO2003002722-A2.XX
PD 09-JAN-2003.

XX
PF 01-JUL-2002; 2002WO-US020890.

XX
PR 29-JUN-2001; 2001US-0302344P.

XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;

XX
DR WPI; 2003-210268/20.

DR N-PSDB; ABZ68333.
XX

PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and

PT Mucin domain gene sequences, useful for treating cancer or asthma,

PT allergv, eczema or autoimmune disease.

PT

PT

PT

XX
PS Claim 10; Page 82; 94pp; English.

The present sequence is a human T cell immunoglobulin domain and mucin domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2, TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with conserved IgV and mucin domains. The locus comprising the TIM family is genetically associated with immune dysfunction, including asthma. The TIM gene family is located within a region of human chromosome 5 that is commonly deleted in malignancies and myelodysplastic syndrome. Variants of TIM-1 and TIM-3 are associated with susceptibility to airway hyperreactivity and allergic T cell responses, and other variants associated with protection against these responses. T cells express TIM proteins, which critically regulate CD4 T cell differentiation. Th1 cells preferentially express TIM-3, while Th2 cells preferentially express TIM-1. TIM polypeptides and polynucleotides are useful for treating cancer, asthma, allergies, eczema or autoimmune diseases.

XX
SQ Sequence 359 AA;

Query Match	100.0%;	Score 275;	DB 6;	Length 359;
Best Local Similarity	100.0%;	Pred. No. 8e-26;		

Query Match	100.0%;	Score 275;	DB 6;	Length 359;
Best Local Similarity	100.0%;	Pred. No. 8e-26;		

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 52
|||||
Db 249 PLYSYTTDGNVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 300

RESULT 8
ABP70438
ID ABP70438 standard; protein; 359 AA.
AC ABP70438;
XX
XX
DT 22-APR-2003 (first entry)
XX
XX Amino acid sequence of human TIM-1 allele 1.
XX
XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
XX WO2003002722-A2.
PN
XX
XX 09-JAN-2003.
PD
PF 01-JUL-2002; 2002WO-US020890.
XX
XX 29-JUN-2001; 2001US-0302344P.
PR
XX (STRD) UNIV LELAND STANFORD JUNIOR.
PA
PI McIntire JJ, Dekruff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX
XX WPI; 2003-210268/20.
DR N-PSDB; ABZ68332.
XX
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
XX
XX Claim 10; Page 80-81; 94pp; English.
PS
XX The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IgV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 275; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 8e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 52
|||||
Db 249 PLYSYTTDGNVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 300

RESULT 9
ABR48174

ID ABR48174 standard; protein; 359 AA.
XX
XX ABR48174;
AC
XX 12-JUN-2003 (first entry)
DT
XX
XX Human bladder cancer associated protein sequence SEQ ID NO:64.
DE
XX
XX Human bladder cancer; cytostatic; gene therapy; vaccine.
KW
XX
OS Homo sapiens.
XX
XX WO2003003906-A2.
PN
XX
XX 16-JAN-2003.
PD
XX
XX 03-JUL-2002; 2002WO-US021338.
PF
XX
XX 03-JUL-2001; 2001US-0302814P.
PR
XX 03-AUG-2001; 2001US-0310099P.
PR
XX 08-NOV-2001; 2001US-0343705P.
PR
XX 13-NOV-2001; 2001US-0350666P.
PR
XX 12-APR-2002; 2002US-0372246P.
PR
XX (EOSE-) EOS BIOTECHNOLOGY INC.
PA
XX
XX Mack DH, Aziz N;
PI
XX
XX WPI; 2003-201532/19.
DR N-PSDB; ACC50985.
XX
XX Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
XX Claim 10; Page 252; 307pp; English.
PS
XX
XX The present invention describes a method for detecting a bladder cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with a polynucleotide
CC that selectively hybridizes to a sequence that is 80 % identical to a
CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC encode the human bladder cancer-associated proteins given in ABR48146 to
CC ABR48242). Bladder cancer-associated sequences from the present invention
CC have cytostatic activities, and can be used in antisense gene therapy and
CC in vaccine production. The method can be used for detecting a bladder
CC cancer-associated transcript in a cell from a patient. The method is
CC useful in diagnosing or treating bladder cancer and in screening for
CC compounds that modulate bladder cancer, such as hormones or antibodies.
CC The nucleic acid molecules from the present invention may be used in
CC various screening and diagnostic methods, and for gene therapy, vaccine
CC and/or antisense/inhibition applications
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 275; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 8e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 52
|||||
Db 249 PLYSYTTDGNVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 300

RESULT 10
AAO26680
ID AAO26680 standard; protein; 359 AA.
XX
XX AAO26680;
AC
XX
XX 20-MAR-2003 (first entry)
DT
XX
XX KIM-1 related protein, SEQ ID No 8.
DE

XX Cytostatic; gene therapy; antigen; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX Homo sapiens.
XX WO200298920-A1.
XX PD 12-DEC-2002.
XX PF 31-MAY-2002; 2002WO-US017402.
XX PR 01-JUN-2001; 2001US-0295449P.
XX PR 04-JUN-2001; 2001US-0295907P.
XX PA (BIOJ) BIOGEN INC.
XX PA (GEHO) GEN HOSPITAL CORP.
XX PI Bailly V, Bonventre J;
XX WPI; 2003-156845/15.
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX Disclosure; Page 40-41; 42pp; English.
XX The invention relates to a novel antibody, antibody derivative or antigen
CC binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a protein relating to the KIM-1
CC protein of the invention
XX
SQ Sequence 359 AA;
Query Match 100.0%; Score 275; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 8e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLYSYTGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
Db 249 PLYSYTGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
RESULT 11
ADE36594
ID ADE36594 standard; protein; 359 AA.
XX ADE36594;
XX 29-JAN-2004 (first entry)
XX Human NOV1b protein SEQ ID NO:4.
XX human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.
XX Homo sapiens.
XX WO2003080856-A2.
XX PD 02-OCT-2003.
XX PF 19-MAR-2003; 2003WO-US008490.
XX PR 19-MAR-2002; 2002US-0365491P.
XX PR 13-SEP-2002; 2002US-0410618P.

XX (CURA-) CURAGEN CORP.
XX Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;
XX WPI; 2003-876927/81.
XX N-PSDB; ADE36593.
XX New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
XX Claim 1; SEQ ID NO 4; 239pp; English.
XX The present invention describes an isolated human NOVX polypeptide, where
CC x is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to it, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOV1b from the present invention.
XX
SQ Sequence 359 AA;
Query Match 100.0%; Score 275; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 8e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 PLYSYTGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
Db 249 PLYSYTGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
RESULT 12
ADN38984
ID ADN38984 standard; protein; 359 AA.
XX ADN38984;
XX 17-JUN-2004 (first entry)
XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:302.
XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX Homo sapiens.
XX WO2003042661-A2.
XX 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-0368809P.
 PR 04-APR-2002; 2002US-0370110P.
 PR 12-APR-2002; 2002US-0372246P.
 PR 05-JUN-2002; 2002US-0386614P.
 PR 16-JUL-2002; 2002US-0396839P.
 PR 22-JUL-2002; 2002US-0397775P.
 PR 22-JUL-2002; 2002US-0397845P.
 PR 09-SEP-2002; 2002US-0409450P.
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Aziz N, Gineburg WM, Gish KC, Glynne R, Hevezl PA;
 PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 XX
 XX WPI; 2003-468649/44.
 DR N-PSDB; ADN38983.
 XX
 XX Determining the presence or absence of a pathological cell in a patient,
 PT useful for diagnosing, prognosing or treating cancer, comprises detecting
 PT a nucleic acid in a biological sample.
 XX
 XX Claim 12; SEQ ID NO 302; 1385pp; English.
 PS
 XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a polypeptide of the invention.
 XX
 XX Sequence 359 AA;
 SQ
 Query Match 100.0%; Score 275; DB 7; Length 359;
 Best Local Similarity 100.0%; Pred. No. 8e-26;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PLYSVTTDGNVTWESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
 DB 249 PLYSVTTDGNVTWESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
 RESULT 13
 ADQ76690
 ID ADQ76690 standard; protein; 359 AA.
 XX
 AC ADQ76690;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Human kidney injury molecule-1 (KIM-1).
 XX
 KW Human; KIM-1; kidney injury molecule-1; immunosuppressive; protozoacide;
 KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
 KW antiallergic; antiinflammatory; vasotropic; nephropic.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Key 1..20
 FT Peptide /label= Signal_peptide
 FT Modified-site 65..67
 FT /note= "N-glycosylated"

FT Modified-site 258..260
 FT /note= "N-glycosylated"
 FT Modified-site 272..274
 FT /note= "N-glycosylated"
 FT Modified-site 286..288
 FT /note= "N-glycosylated"
 FT Domain 290..311
 FT /note= "Transmembrane domain"
 FT Region 324..359
 FT /note= "C-terminal region unique to this splice variant"
 XX
 XX WO2004060041-A2.
 PN 22-JUL-2004.
 XX
 XX 29-DEC-2003; 2003WO-US041294.
 XX
 XX 30-DEC-2002; 2002US-0436934P.
 PR (BIOG-) BIOGEN IDEC MA INC.
 PA Remmert PD;
 XX
 XX WPI; 2004-534277/51.
 DR
 XX Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
 PT Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
 PT activation, or inhibiting production of a subset of antibodies against
 PT one or more antigens.
 XX
 PS Disclosure; SEQ ID NO 1; 68pp; English.
 XX
 CC The present sequence is that of human kidney injury molecule-1 (KIM-1), a
 CC type I cell membrane glycoprotein and member of the T cell immunoglobulin
 CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids
 CC ADQ76696 or 359 amino acids, depending on splice variation. It has been
 CC discovered that treatment of a mammal with a KIM-1 antagonist alters the
 CC interaction of T cells and other immune system cells, e.g. dendritic
 CC cells, monocytes, macrophages and B cells, and thereby strongly
 CC suppresses an IgG response to an antigen. Such treatment also eliminates
 CC IgG1 production by memory B cells in response to subsequent challenge
 CC with the antigen. Blockage of the binding of KIM-1 to its receptor
 CC reduces secretion of interferon-gamma by immune cells engaged in an
 CC antigen response in the mixed lymphocyte response assay. Based on these
 CC discoveries, the invention provides methods for therapeutically
 CC modulating immune function in autoimmune diseases and other disorders of
 CC the mammalian immune system. A method is claimed for inhibiting
 CC signalling between a T cell and a second cell, e.g. an antigen-presenting
 CC cell, in a mammal. The method involves identifying a mammal with an
 CC immune disease or disorder, or one preparing to receive a tissue graft,
 CC and administering a KIM-1 antagonist, especially a polypeptide
 CC comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a
 CC KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding
 CC fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a
 CC soluble polypeptide, which can include a KIM-1 mucin domain in addition
 CC to the KIM Ig domain, and may include a heterologous moiety such as an Fc
 CC moiety. The antagonist may be conjugated to a polymer such as
 CC polyethylene glycol. The KIM-1 antagonist is used in methods of:
 CC inhibiting activation of a B cell in a mammal; inhibiting disease relapse
 CC in an autoimmune disease; inhibiting epitope spreading in an autoimmune
 CC disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis,
 CC autoimmune haemolytic anaemia, Chagas disease, Grave's disease,
 CC idiopathic thrombocytopenia purpura, Wegener's granulomatosis,
 CC polyarteritis nodosa, rapidly progressive crescentic glomerulonephritis,
 CC graft-versus-host disease, or systemic lupus nephritis) or inflammatory
 CC disease or disorder (e.g. inflammatory bowel diseases such as Crohn's
 CC disease, ulcerative colitis, and ileitis); and inhibiting secretion of
 CC IFN-gamma by lymphocytes.
 XX
 XX Sequence 359 AA;
 SQ
 Query Match 100.0%; Score 275; DB 8; Length 359;
 Best Local Similarity 100.0%; Pred. No. 8e-26;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 52
AC
XX
DT 22-APR-2003 (first entry)
XX
DE Amino acid sequence of human TIM-1 allele 1
XX
KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 208 /note= "encoded by RCG"
FT
XX
PN WO2003002722-A2.
XX
PD 09-JAN-2003.
XX
PF 01-JUL-2002; 2002WO-US020890.
XX
PR 29-JUN-2001; 2001US-0302344P.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX
DR WPI; 2003-210268/20.
DR N-PSDB; ABZ68336.
XX
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
XX
PS Claim 10; Page 86-87; 94pp; English.
XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IGV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
SQ Sequence 364 AA;

Query Match 100.0%; Score 275; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.1e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 52
DB 254 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 305

RESULT 14
ABP70442
ID ABP70442 standard; protein; 364 AA.
XX
AC ABP70442;
XX
DT 22-APR-2003 (first entry)
XX
DE Amino acid sequence of human TIM-1 allele 5.
XX
KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
PN WO2003002722-A2.
XX
PD 09-JAN-2003.
XX
PF 01-JUL-2002; 2002WO-US020890.
XX
PR 29-JUN-2001; 2001US-0302344P.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX
DR WPI; 2003-210268/20.
DR N-PSDB; ABZ68336.
XX
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
XX
PS Claim 10; Page 86-87; 94pp; English.
XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IGV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
SQ Sequence 364 AA;

Query Match 100.0%; Score 275; DB 6; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.1e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 52
DB 254 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 305

RESULT 15
ABP70440

ID ABP70440 standard; protein; 365 AA.
XX
AC ABP70440;
XX
DT 22-APR-2003 (first entry)
XX
DE Amino acid sequence of human TIM-1 allele 3.
XX
KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 208 /note= "encoded by RCG"
FT
XX
PN WO2003002722-A2.
XX
PD 09-JAN-2003.
XX
PF 01-JUL-2002; 2002WO-US020890.
XX
PR 29-JUN-2001; 2001US-0302344P.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX
DR WPI; 2003-210268/20.
DR N-PSDB; ABZ68334.
XX
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
XX
PS Claim 10; Page 83-84; 94pp; English.
XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IGV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
SQ Sequence 365 AA;

Query Match 100.0%; Score 275; DB 6; Length 365;
Best Local Similarity 100.0%; Pred. No. 8.2e-26;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 52
DB 255 PLYSYTTDGNVTVESSDGLWNNQTLFLEHSLTANTTKGIYAGVCISVL 306

Search completed: June 29, 2005, 09:09:09
Job time : 37.9823 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 7.15128 Seconds
(without alignments)
542.805 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300

Perfect score: 275

Sequence: 1 PLYSYTGDNDVTESDGL.....SLLTANTTKGIYAGVCISVL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA*

- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	275	100.0	334	4	US-09-197-970B-7
2	228	82.9	451	1	US-08-287-001A-2
3	228	82.9	451	5	PCT-US95-09941-2
4	78	28.4	307	4	US-09-197-970B-3
5	62	22.5	606	4	US-09-603-208A-56
6	60	21.8	602	1	US-08-168-091A-2
7	60	21.8	602	1	US-08-428-926-5
8	60	21.8	602	1	US-08-428-927-5
9	60	21.8	602	1	US-08-428-298-5
10	60	21.8	602	1	US-08-339-517-5
11	60	21.8	1070	3	US-08-697-954-2
12	60	21.8	1260	4	US-09-328-352-6746
13	58	21.1	981	4	US-09-991-258-13
14	56.5	20.5	1039	4	US-09-270-767-41514
15	55.5	20.2	137	4	US-09-540-236-2771
16	55.5	20.2	495	4	US-09-691-270A-24
17	55	20.0	386	4	US-09-328-352-8185
18	55	20.0	463	4	US-09-375-140-10
19	55	20.0	548	4	US-09-375-140-11
20	54.5	19.8	323	6	5260223-4
21	54.5	19.8	323	6	5260223-4
22	54.5	19.8	1619	4	US-09-328-352-7347
23	54	19.6	357	4	US-09-248-796A-21669
24	53.5	19.5	72	4	US-09-489-847-366
25	53.5	19.5	73	4	US-09-489-847-241
26	53.5	19.5	224	4	US-09-248-796A-18758
27	53.5	19.5	10182	3	US-09-134-001C-3159

28	53	19.3	112	3	US-09-134-001C-3352	Sequence 3352, Ap
29	53	19.3	190	3	US-08-341-018-44	Sequence 44, Appl
30	53	19.3	190	3	US-08-470-335-217	Sequence 217, App
31	53	19.3	190	3	US-08-470-339-217	Sequence 217, App
32	53	19.3	190	4	US-08-467-602-411	Sequence 411, App
33	53	19.3	210	4	US-08-467-602-191	Sequence 191, App
34	53	19.3	210	4	US-08-411-295F-315	Sequence 315, App
35	53	19.3	213	4	US-08-467-603-186	Sequence 186, App
36	53	19.3	213	4	US-08-411-295F-310	Sequence 310, App
37	53	19.3	219	4	US-08-467-602-200	Sequence 200, App
38	53	19.3	219	4	US-08-411-295F-324	Sequence 324, App
39	53	19.3	222	4	US-08-467-602-187	Sequence 187, App
40	53	19.3	222	4	US-08-411-295F-321	Sequence 321, App
41	53	19.3	233	4	US-08-467-603-194	Sequence 194, App
42	53	19.3	233	4	US-08-411-295F-318	Sequence 318, App
43	53	19.3	242	4	US-08-467-602-203	Sequence 203, App
44	53	19.3	242	4	US-08-411-295F-327	Sequence 327, App
45	53	19.3	309	4	US-09-248-796A-25339	Sequence 25339, A

ALIGNMENTS

RESULT 1

US-09-197-970B-7
; Sequence 7, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Biogen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09197,970B
FILING DATE: 23-No. 6664385-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,228
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Levine, Leslie M.
REGISTRATION NUMBER: 35,245
REFERENCE/DOCKET NUMBER: A010 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 679-2810
TELEFAX: (617) 679-2838
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 334 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-197-970B-7

Query Match 100.0%; Score 275; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.2e-28;

```
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
    |||||||
Db 249 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300
    |||||||

RESULT 2
US-08-287-001A-2
; Sequence 2, Application US/08287001A
; Patent No. 5622861
; GENERAL INFORMATION:
; APPLICANT: KAPLAN, GERARDO
; APPLICANT: FEINSTONE, STEPHEN M.
; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
; TITLE OF INVENTION: OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,001A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US-287-001A-2

Query Match 82.9%; Score 228; DB 1; Length 451;
Best Local Similarity 84.6%; Pred. No. 3.5e-22;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
    |||||||
Db 329 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 380
    |||||||

RESULT 3
PCT-US95-09941-2
; Sequence 2, Application PC/TUS9509941
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
; TITLE OF INVENTION: OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA

Query Match 82.9%; Score 228; DB 1; Length 451;
Best Local Similarity 84.6%; Pred. No. 3.5e-22;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
    |||||||
Db 329 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 380
    |||||||

RESULT 4
US-09-197-970B-3
; Sequence 3, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-NO. 6664385-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
```

```
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09941
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/287,001
; FILING DATE: 5 AUG 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09941-2

Query Match 82.9%; Score 228; DB 5; Length 451;
Best Local Similarity 84.6%; Pred. No. 3.5e-22;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
    |||||||
Db 329 PLYSYTGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 380
    |||||||

RESULT 4
US-09-197-970B-3
; Sequence 3, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-NO. 6664385-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
```

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 679-2810
TELEFAX: (617) 679-2838
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 307 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-197-9708-3

Query Match 28.4%; Score 78; DB 4; Length 307;
Best Local Similarity 42.0%; Pred. No. 0.026;
Matches 21; Conservative 7; Mismatches 18; Indels 4; Gaps 3;

QY 4 SYT-TDGNDFVTSSDGLMNNOTQFLFLEHSLLTANTTKGIYAGVCISVL 52
DB 199 SYTPADWNGT-VT-SSEAWNNHTVRIPLRKP--QRNPTKGFVGMVAAL 245

RESULT 5

US-09-603-208A-56
Sequence 56, Application US/09603208A
Patent No. 6822084
GENERAL INFORMATION:
APPLICANT: Pompejus, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberhauer, Gregor
APPLICANT: Lee, Heung-Shick
APPLICANT: Kim, Hyung-Joon

TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS

FILE REFERENCE: BGI-124CP
CURRENT APPLICATION NUMBER: US/09/603,208A
CURRENT FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: 60/142692
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: 60/151214
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19930429.7
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: DE 19931413.6
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931457.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19931541.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932209.0
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932230.9
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932914.1
PRIOR FILING DATE: 1999-07-14
PRIOR APPLICATION NUMBER: DE 19940764.9
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: DE 19941382.7
PRIOR FILING DATE: 1999-08-31
NUMBER OF SEQ ID NOS: 306
SEQ ID NO 56
LENGTH: 606
TYPE: PRT
ORGANISM: Corynebacterium glutamicum

US-09-603-208A-56

Query Match 22.5%; Score 62; DB 4; Length 606;
Best Local Similarity 32.6%; Pred. No. 8.7;
Matches 14; Conservative 5; Mismatches 24; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNDFVTSSDGLMNNOTQFLFLEHSLLTANTTKGI 43
DB 559 PLVEYQKEGDMFNMGKDIKEETVRQLFSLSSSSSKTRKSL 601

RESULT 6

US-08-168-091A-2
Sequence 2, Application US/08168091A
Patent No. 5665862
GENERAL INFORMATION:
APPLICANT: Fischbach, Gerald.
APPLICANT: Falls, Douglas R.
APPLICANT: Rosen, Kenneth M.
APPLICANT: Corfaa, Gabriel
TITLE OF INVENTION: Neurotrophic Factor
NUMBER OF SEQUENCES: 47
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE AND COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/168,091A
FILING DATE: 15-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/953,742
FILING DATE: 29-SEP-1992

ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: HMI-002CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 602 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-168-091A-2

Query Match 21.8%; Score 60; DB 1; Length 602;
Best Local Similarity 31.7%; Pred. No. 16;
Matches 13; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 3 YSYTTDGNDFVTSSDGLMNNOTQFLF--EHSLLTANTTK 41
DB 301 YTSITTHSMVTVTQTPSHSNGHTSILSESHSVLVSSSVE 341

RESULT 7

US-08-428-926-5
Sequence 5, Application US/08428926
Patent No. 5667780
GENERAL INFORMATION:

APPLICANT: Ho, Wei-Hsien
APPLICANT: Osheroif, Phyllis L.
TITLE OF INVENTION: SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 460 Point San Bruno Blvd
CITY: South San Francisco
STATE: California

ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 46839-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1070 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-697-954-2

Query Match 21.8%; Score 60; DB 3; Length 1070;
Best Local Similarity 31.7%; Pred. No. 33;
Matches 13; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 3 YSYTTDGNVTVESDGLWNNNQTLFL--EHSLLTANTTK 41
||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 586 YTSTTHSMVTQTSPSHWSNGHTESILSESHVLVSSSV 626

RESULT 12

US-09-328-352-6746
Sequence 6746, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
CURRENT APPLICATION NUMBER: GTC99-03PA
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 6746
LENGTH: 1260
TYPE: PR1
ORGANISM: Acinetobacter baumannii

US-09-328-352-6746

Query Match 21.8%; Score 60; DB 4; Length 1260;
Best Local Similarity 32.7%; Pred. No. 41;
Matches 16; Conservative 9; Mismatches 22; Indels 2; Gaps 1;

QY 3 YSYTTDGNVTVESDGLWNNNQTLFLHSLLTANTTKGIYAGVCISV 51
||| : ||| : ||| : ||| : ||| : ||| : ||| :
DB 228 YSYDMLGEAALTDDHABRYFNDTQAI--HAIGKASNGKVVDGPGISI 274

RESULT 13

US-09-991-258-13
Sequence 13, Application US/09991258
Patent No. 6783939
GENERAL INFORMATION:
APPLICANT: Olmsted, Robert
APPLICANT: Keith, Paula
APPLICANT: Dryga, Sergey
APPLICANT: Caley, Ian
APPLICANT: Maughan, Maureen
APPLICANT: Johnston, Robert
APPLICANT: Davis, Nancy
APPLICANT: Swanson, Ronald
TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOSES WITH MODIFIED HIV GENES FOR USE
FILE REFERENCE: VACCINES
CURRENT APPLICATION NUMBER: 01113.0001U3
CURRENT FILING DATE: 2001-11-16
PRIOR APPLICATION NUMBER: 09/902,537
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: 60/216,995
PRIOR FILING DATE: 2000-07-07

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:58:20 ; Search time 32.5894 Seconds
(without alignments)
613.589 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300

Perfect score: 275

Sequence: 1 PLYSYTDCGNTVTSSDGL.....SLLTANTTKGIYAGVCISVL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/ECT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/ECTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10F_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	275	100.0	334	17	US-10-655-506-7
2	275	100.0	334	17	US-10-718-321-7
3	275	100.0	339	17	US-10-391-939A-2
4	275	100.0	339	17	US-10-391-939A-28
5	275	100.0	359	14	US-10-188-012-17
6	275	100.0	359	14	US-10-188-012-19
7	275	100.0	359	15	US-10-295-027-302
8	275	100.0	359	15	US-10-188-832-64
9	275	100.0	359	17	US-10-391-939A-4
10	275	100.0	359	17	US-10-663-497-17
11	275	100.0	359	17	US-10-663-497-19

12	275	100.0	359	17	US-10-718-321-8	Sequence 8, Appli
13	275	100.0	359	17	US-10-847-918-25	Sequence 25, Appl
14	275	100.0	364	14	US-10-188-012-25	Sequence 25, Appl
15	275	100.0	364	17	US-10-663-497-25	Sequence 25, Appl
16	275	100.0	365	14	US-10-188-012-21	Sequence 21, Appl
17	275	100.0	365	17	US-10-663-497-21	Sequence 21, Appl
18	270	98.2	359	14	US-10-188-012-23	Sequence 23, Appl
19	270	98.2	359	17	US-10-663-497-23	Sequence 23, Appl
20	270	98.2	364	14	US-10-188-012-27	Sequence 27, Appl
21	270	98.2	364	17	US-10-663-497-27	Sequence 27, Appl
22	225	81.8	81	17	US-10-718-321-6	Sequence 6, Appli
23	189	68.7	263	17	US-10-391-939A-6	Sequence 6, Appli
24	189	68.7	263	17	US-10-391-939A-32	Sequence 32, Appl
25	189	68.7	263	17	US-10-391-939A-39	Sequence 39, Appl
26	189	68.7	263	17	US-10-805-177-50	Sequence 50, Appl
27	99	36.0	18	17	US-10-718-321-1	Sequence 1, Appli
28	79	28.7	305	14	US-10-188-012-1	Sequence 1, Appli
29	79	28.7	305	17	US-10-663-497-1	Sequence 1, Appli
30	78	28.4	307	17	US-10-655-506-3	Sequence 3, Appli
31	76.5	27.8	282	14	US-10-188-012-3	Sequence 3, Appli
32	76.5	27.8	282	17	US-10-663-497-3	Sequence 3, Appli
33	66	24.0	666	16	US-10-437-963-196644	Sequence 196644,
34	65	23.6	554	14	US-10-125-692-21	Sequence 21, Appl
35	60.5	22.0	727	16	US-10-296-723A-14	Sequence 14, Appl
36	60	21.8	244	15	US-10-424-599-239255	Sequence 239255,
37	60	21.8	285	15	US-10-424-599-255038	Sequence 255038,
38	60	21.8	338	15	US-10-425-114-37036	Sequence 37036, A
39	60	21.8	338	15	US-10-412-699B-514	Sequence 514, App
40	59.5	21.6	125	15	US-10-424-599-220078	Sequence 220078,
41	58.5	21.6	305	14	US-10-188-012-5	Sequence 5, Appli
42	59.5	21.6	305	14	US-10-188-012-7	Sequence 7, Appli
43	59.5	21.6	305	17	US-10-663-497-5	Sequence 5, Appli
44	59.5	21.6	305	17	US-10-663-497-7	Sequence 7, Appli
45	58	21.1	272	16	US-10-425-115-288232	Sequence 288232,

ALIGNMENTS

RESULT 1

US-10-655-506-7
; Sequence 7, Application US/10655506
; Publication No. US20050089868A1
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/655,506
; FILING DATE: 04-Sep-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970
; FILING DATE: 23-Nov-1998
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-655-506-7

Query Match          100.0%; Score 275; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 2
US-10-718-321-7
; Sequence 7, Application US/10718321
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-7

Query Match          100.0%; Score 275; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 3
US-10-391-939A-2
; Sequence 2, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Giot, Loic;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Mesri, Mehdi;
; APPLICANT: Ooi, Chean Eng;
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-655-506-7

Query Match          100.0%; Score 275; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 4
US-10-391-939A-28
; Sequence 28, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Giot, Loic;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Mesri, Mehdi;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glennnda;
; APPLICANT: Starling, Gary;
; APPLICANT: Tse, Kam-Fai
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-048
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 28
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-28

Query Match          100.0%; Score 275; DB 17; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 229 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 280

RESULT 5
US-10-188-012-17
; Sequence 17, Application US/10188012
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; TITLE OF INVENTION: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-188-012-17

Query Match          100.0%; Score 275; DB 17; Length 339;
Best Local Similarity 100.0%; Pred. No. 4.2e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 229 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 280
```


; Sequence 17, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-188-012-17

Query Match 100.0%; Score 275; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
|||||
DB 249 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 6

US-10-188-012-19
; Sequence 19, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-188-012-19

Query Match 100.0%; Score 275; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
|||||
DB 249 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 7

US-10-295-027-302
; Sequence 302, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 302
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-302

Query Match 100.0%; Score 275; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
|||||
DB 249 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 8

US-10-188-832-64
; Sequence 64, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Nataasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22

```
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-64

Query Match      100.0%; Score 275; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
   |||||||
Db 249 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 9
US-10-391-939A-4
; Sequence 4, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Guot, Loic ;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Meari, Mehdi ;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca ;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glenda ;
; APPLICANT: Starling, Gary ;
; APPLICANT: Tse, Kam-Fai
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-048
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraSeqIst version 0.1
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-4

Query Match      100.0%; Score 275; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
   |||||||
Db 249 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 10
US-10-663-497-17
; Sequence 17, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-663-497-17

Query Match      100.0%; Score 275; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
   |||||||
Db 249 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 11
US-10-663-497-19
; Sequence 19, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; TITLE OF INVENTION: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-663-497-19

Query Match      100.0%; Score 275; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
   |||||||
Db 249 PLSYTTDGDNDVTVESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
```

```
RESULT 12
US-10-718-321-8
; Sequence 8, Application US/10718321
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; TITLE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-8

Query Match      100.0%; Score 275; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
DB 249 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 13
US-10-847-918-25
; Sequence 25, Application US/10847918
; Publication No. US20050119210A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Be, Xiaobing
; APPLICANT: Liu, Wei
; APPLICANT: Slonim, Donna
; APPLICANT: Howes, Steve
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating Cancers
; FILE REFERENCE: 031896-026000 (AM101264)
; CURRENT APPLICATION NUMBER: US/10/847,918
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US 60/471,729
; PRIOR FILING DATE: 2003-05-20
; NUMBER OF SEQ ID NOS: 14937
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-847-918-25

Query Match      100.0%; Score 275; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
DB 249 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

RESULT 14
US-10-188-012-25
; Sequence 25, Application US/10188012
```

```
Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 364
; TYPE: PRT
; ORGANISM: H. sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(364)
; OTHER INFORMATION: TIM-1 allele 5
US-10-188-012-25

Query Match      100.0%; Score 275; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
DB 254 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 305

RESULT 15
US-10-663-497-25
; Sequence 25, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; TITLE OF INVENTION: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 364
; TYPE: PRT
; ORGANISM: H. sapiens
; NAME/KEY: VARIANT
; LOCATION: (1)...(364)
; OTHER INFORMATION: TIM-1 allele 5
US-10-663-497-25

Query Match      100.0%; Score 275; DB 17; Length 364;
Best Local Similarity 100.0%; Pred. No. 4.6e-28;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
DB 254 PLYSYTTDGNVTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 305
```

Search completed: June 29, 2005, 09:33:02
Job time : 33.5894 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:46:27 ; Search time 202.106 Seconds
(without alignments)
635.503 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLHLADSVAG.....FPKKEVQQLRPHKSCIHORE 334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 487488

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99	5.7	18	17	US-10-718-321-1
2	98	5.6	35	15	US-10-434-269-23
3	81	4.6	41	15	US-10-057-783A-37
4	79	4.5	14	17	US-10-805-177-87
5	75	4.3	35	15	US-10-434-269-21
6	73	4.2	45	14	US-10-228-063-53
7	70.5	4.0	43	9	US-09-864-761-44813
8	70	4.0	12	17	US-10-805-177-88
9	67	3.8	12	17	US-10-805-177-92
10	67	3.8	34	9	US-09-916-494A-24
11	67	3.8	38	17	US-10-933-404-5
					Sequence 1, Appli
					Sequence 23, Appl
					Sequence 37, Appl
					Sequence 87, Appl
					Sequence 21, Appl
					Sequence 53, Appl
					Sequence 44813, A
					Sequence 88, Appl
					Sequence 92, Appl
					Sequence 24, Appl
					Sequence 5, Appli

12	66.5	3.8	37	17	US-10-935-254-45	Sequence 45, Appl
13	65	3.7	47	16	US-10-425-115-354965	Sequence 354965,
14	64	3.7	39	17	US-10-622-893A-10	Sequence 10, Appl
15	63	3.6	32	16	US-10-327-598-400	Sequence 400, App
16	63	3.6	47	13	US-10-105-934-17	Sequence 17, Appl
17	63	3.6	47	17	US-10-895-676-17	Sequence 17, Appl
18	62.5	3.6	40	16	US-10-425-115-352983	Sequence 352983,
19	62	3.5	34	9	US-09-864-761-37295	Sequence 37295, A
20	62	3.5	47	16	US-10-437-963-203485	Sequence 203485,
21	61.5	3.5	30	15	US-10-434-269-22	Sequence 22, Appl
22	61	3.5	34	17	US-10-933-404-6	Sequence 6, Appli
23	60.5	3.5	46	16	US-10-425-115-209986	Sequence 209986,
24	60	3.4	50	9	US-09-864-761-40317	Sequence 40317, A
25	59.5	3.4	45	9	US-09-864-761-39258	Sequence 39258, A
26	59.5	3.4	46	9	US-09-864-761-41678	Sequence 41678, A
27	59.5	3.4	48	9	US-09-799-514-18	Sequence 18, Appl
28	59	3.4	10	17	US-10-805-177-89	Sequence 89, Appl
29	59	3.4	15	17	US-10-718-321-5	Sequence 5, Appli
30	59	3.4	49	16	US-10-425-115-288406	Sequence 288406,
31	58	3.3	47	9	US-09-731-449-22	Sequence 22, Appl
32	58	3.3	47	9	US-09-731-449-23	Sequence 23, Appl
33	58	3.3	47	14	US-10-254-426-22	Sequence 22, Appl
34	58	3.3	47	14	US-10-254-426-23	Sequence 23, Appl
35	57	3.3	33	15	US-10-424-599-324859	Sequence 224859,
36	57	3.3	43	14	US-10-428-662-109	Sequence 109, App
37	57	3.3	44	15	US-10-424-599-189797	Sequence 189797,
38	57	3.3	45	15	US-10-424-599-258666	Sequence 258666,
39	56	3.2	10	17	US-10-805-177-93	Sequence 93, Appl
40	56	3.2	38	17	US-10-877-849-26	Sequence 26, Appl
41	56	3.2	45	16	US-10-425-115-288694	Sequence 288694,
42	56	3.2	47	16	US-10-425-115-195346	Sequence 195346,
43	56	3.2	50	17	US-10-492-403A-2	Sequence 2, Appli
44	55.5	3.2	40	10	US-09-798-889-84	Sequence 84, Appl
45	55.5	3.2	40	15	US-10-633-680-84	Sequence 84, Appl

ALIGNMENTS

RESULT 1
US-10-718-321-1
; Sequence 1, Application US/10718321
; Publication NO. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10718,321
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-1

Query Match 5.7%; Score 99; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 264 SSDGLWNNQTLFLEHS 281
|||||

Db 1 SSDGLWNNNOTOLEHS 18

RESULT 2

```

US-10-434-269-23
; Sequence 23, Application US/10434269
; Publication No. US20040053296A1
; GENERAL INFORMATION:
; APPLICANT: ITO, Makoto, et al.
; TITLE OF INVENTION: A METHOD FOR TAILORING
; FILE REFERENCE: 1492-0588P
; CURRENT APPLICATION NUMBER: US/10/434
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 10/434
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Insect intestinal mucin
US-10-434-269-23

```

Query Match 5.6%; Score 98; DB 15; Length 35;
Best Local Similarity 59.0%; Pred. No. 6.2;
Matches 23; Conservative 0; Mismatches 10; Indels 6; Gaps 2;

Qy 152 TTTTTPPTTT - VPTTMSIPPTTTTTPPTTMTVSTTTTSVPPTTT 189
 |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
 db 2 TTTTCAPTTTTCAPTTTTCAPTTTTCAPTTTTCAPTTTTCAPTTTTCAPTTTT 35
 -----TTTTCAPTTTT

RESULT 3

```

RESULTS 3
US-10-057-783A-37
; Sequence 37, Application US/10057783A
; Publication No. US20040091955A1
; GENERAL INFORMATION:
; APPLICANT: Forster, Anthony C.
; TITLE OF INVENTION: Process and compo
; TITLE OF INVENTION: peptidomimetic
; FILE REFERENCE: 1
; CURRENT APPLICATION NUMBER: US/10/05
; CURRENT FILING DATE: 2002-01-25
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 37
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of A
; -OTHER INFORMATION: DNA
US-10-057-783A-37

```

Query Match 4.6%; Score 81; DB 15; Length 41;
Best Local Similarity 49.0%; Pred. No. 1.1e+02;
Matches 24; Conservative 2; Mismatches 13; Indels 10; Gaps 2;

Qy 129 VTTTPVITVPVTTRTSITVPITTTTVPPTTMSPTTTTVPITM 177
|||:||||||| ||| ||| ||| |

pB 2 VTWTWTTTWTVW---TTWTTWTTWTTW-----TWTTWTTV 40

RESULT 4

US-10-805-177-87
; Sequence 87, Application US/10805177
; Publication No. US2005008449A1
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Chen, Francine
; APPLICANT: Bezabeh, Binyam
; APPLICANT: Foltz, Ian
; APPLICANT: Tse, Kam Fai

```

? APPLICANT: Jeffers, Michael
? APPLICANT: Meeri, Mehdi
? APPLICANT: Starling, Gary
? APPLICANT: Mezes, Peter
? APPLICANT: Khramtsov, Nikolai
? TITLE OF INVENTION: ANTIBODIES AGAINST T CELL IMMUNOGLOBULIN
? TITLE OF INVENTION: DOMAIN AND MUCIN DOMAIN 1 (TIM-1) ANTIGEN AND USES THEREOF
? FILE REFERENCE: ABXCUR.006A
? CURRENT APPLICATION NUMBER: US/10/805,177
? CURRENT FILING DATE: 2004-03-19
? PRIOR APPLICATION NUMBER: 60/456,652
? PRIOR FILING DATE: 2003-03-19
? NUMBER OF SEQ ID NOS: 141
? SOFTWARE: FastSEQ for Windows Version 4.0
? SEQ ID NO 87
? LENGTH: 14
? TYPE: PRT
? ORGANISM: Homo Sapiens
? US-10-805-177-87

```

```
Query Match      4.5%; Score 79; DB 17; Length 14;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 209 PMPLPRQNHEPVAT 22
Db 1 PMPLPRQNHEPVAT 14

RESULT 5

```

US-10-434-269-21
; Sequence 21, Application US/10434269
; Publication No. US20040053296A1
; GENERAL INFORMATION:
; APPLICANT: ITO, Makoto, et al.
; TITLE OF INVENTION: A METHOD FOR TARGETING A POLYPEPTIDE ONTO CELLULAR SURFACE
; FILE REFERENCE: 1422-0588P
; CURRENT APPLICATION NUMBER: US/10/434,269
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 10/434,269
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-10-434-269-21

```

Query Match 4.3%; Score 75; DB 15; Length 35;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 15; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 174 PTTMTVSTTTSVPPTTTSIPPTTTSVP 198
|| : || || || : || || : || ||
pb 2 PPTNSPTSTPTSTPTSTPTSTPTSTPT 26

RESULT 6

```

US-10-228-063-53
; Sequence 53, Application US/10228063
; Publication No. US20030135885A1
; GENERAL INFORMATION:
; APPLICANT: Lanahan, Mike
; TITLE OF INVENTION: Self-processing Plants and Plant Parts
; FILE REFERENCE: 109846.317
; CURRENT APPLICATION NUMBER: US/10/228,063
; CURRENT FILING DATE: 2002-12-12
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 45
; TYPE: PRT

```


Db 4 STTGPSETTSPALTWTMRP--TSMPLTWQKSYITISKFT 44

RESULT 14
US-10-622-893A-10
; Sequence 10, Application US/10622893A
; Publication No. US20050014935A1
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Sheng
; APPLICANT: Datta, Abhijit
; APPLICANT: Wang, Yuping
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
; FILE REFERENCE: DETERMINATION OF GLYCATED PROTEINS
; CURRENT APPLICATION NUMBER: US/10/622,893A
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 39
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence homology between the N-terminal sequence of Amadoriase I
US-10-622-893A-10

Query Match 3.7%; Score 64; DB 17; Length 39;
Best Local Similarity 61.5%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 6; Indels 4; Gaps 1;
QY 146 TSTVTPTTTTPTTPTTPTTPTTMSIPTTT 171
Db 18 TSTTTTTTTTTTTTTTTT-----TTT 39

RESULT 15
US-10-327-598-400
; Sequence 400, Application US/10327598
; Publication No. US20040181039A1
; GENERAL INFORMATION:
; APPLICANT: Krah, Eugene
; APPLICANT: Guo, Honliang
; APPLICANT: Aiyappa, Ashok
; APPLICANT: Lawton, Robert
; TITLE OF INVENTION: Canine Immunoglobulin Variable Domains, Caninized Antibodies, and
; FILE REFERENCE: 01-799-A
; CURRENT APPLICATION NUMBER: US/10/327,598
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: US 60/344,874
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 1139
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 400
; LENGTH: 32
; TYPE: PRT
; ORGANISM: canis familiaris;
US-10-327-598-400

Query Match 3.6%; Score 63; DB 16; Length 32;
Best Local Similarity 48.0%; Pred. No. 1.5e+03;
Matches 12; Conservative 2; Mismatches 11; Indels 0; Gaps 0;
QY 81 GDLRRDVSLLTIENTAVSDSGVYCC 105
Db 8 GSGSGTDCLTISRVEADDAVGVC 32

Search completed: June 29, 2005, 10:14:19
Job time : 202.106 secs

THIS PAGE BLANK (04/10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:33:11 ; Search time 60.3694 Seconds
(without alignments)
413.004 Million cell updates/sec

Title: US-10-718-321-7

Perfect score: 1749

Sequence: 1 MHPQVILSLHLADSVAG.....PFKKEVQQLRPHKSCIHORE 334

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 272444

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/ACTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	103	5.9	41	4	US-09-060-767B-5
2	86.5	4.9	41	4	US-09-060-767B-8
3	70.5	4.0	29	3	US-09-043-731-23
4	67	3.8	34	2	US-08-169-948B-24
5	67	3.8	34	2	US-08-448-873-24
6	67	3.8	34	3	US-08-382-452D-24
7	67	3.8	34	4	US-08-507-362A-12
8	67	3.8	34	4	US-09-516-494A-24
9	67	3.8	46	3	US-08-856-074A-39
10	64	3.7	25	4	US-09-060-767B-7
11	61.5	3.5	26	2	US-08-288-059-2
12	60.5	3.5	25	4	US-09-060-767B-4
13	59	3.4	40	2	US-08-530-569B-7
14	59	3.4	41	1	US-08-597-495B-28
15	59	3.4	41	3	US-09-068-051A-28
16	58	3.3	37	3	US-08-814-052-37
17	58	3.3	37	3	US-08-812-829-29
18	58	3.3	45	1	US-08-361-920-19
19	58	3.3	45	1	US-08-479-939-19
20	58	3.3	45	1	US-08-483-432-19
21	57	3.3	43	4	US-08-454-899G-109
22	55	3.1	45	1	US-08-451-947-98
23	55	3.1	45	2	US-08-424-826A-98
24	55	3.1	45	3	US-08-928-694-98
25	55	3.1	45	4	US-08-450-842-98
26	55	3.1	45	4	US-08-451-390-98
27	55	3.1	45	5	PCT-US91-06950-98

```

28 54.5 3.1 43 2 US-08-169-948B-20 Sequence 20, Appl
29 54.5 3.1 43 2 US-08-448-873-20 Sequence 20, Appl
30 54.5 3.1 43 3 US-08-382-452D-20 Sequence 20, Appl
31 54.5 3.1 43 4 US-09-916-494A-20 Sequence 20, Appl
32 54.5 3.1 49 4 US-09-461-325-425 Sequence 425, App
33 54.5 3.1 49 4 US-10-012-542-425 Sequence 425, App
34 54.5 3.1 49 4 US-10-115-123-425 Sequence 425, App
35 53.5 3.1 45 4 US-09-270-767-35364 Sequence 35364, A
36 53.5 3.1 45 4 US-09-270-767-50581 Sequence 50581, A
37 53.5 3.1 50 5 PCT-US91-02942-9 Sequence 9, Appl
38 53 3.0 23 2 US-08-833-807-2 Sequence 2, Appl
39 53 3.0 23 2 US-09-223-043-2 Sequence 2, Appl
40 53 3.0 23 4 US-09-593-870A-2 Sequence 2, Appl
41 53 3.0 32 1 US-08-137-117D-132 Sequence 132, App
42 53 3.0 32 2 US-08-436-717-132 Sequence 132, App
43 53 3.0 44 3 US-08-955-937A-13 Sequence 13, Appl
44 53 3.0 44 3 US-09-300-985-13 Sequence 13, Appl
45 53 3.0 45 4 US-09-270-767-57229 Sequence 57229, A

```

ALIGNMENTS

RESULT 1

US-09-060-767B-5
; Sequence 5, Application US/09060767B
; Patent No. 6720152
; GENERAL INFORMATION:
; APPLICANT: Weil, Gary
; APPLICANT: Chandrashekar, Ramaswamy
; TITLE OF INVENTION: Diagnosis of Histoplasmosis Using Antigens Specific for
; FILE OF INVENTION: H. capsulatum
; FILE REFERENCE: BJCH 9986
; CURRENT APPLICATION NUMBER: US/09/060,767B
; CURRENT FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/043,332
; PRIOR FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Leishmania
US-09-060-767B-5

Query Match 5.9%; Score 103; DB 4; Length 41;
Best Local Similarity 56.2%; Pred. No. 0.095; 10; Indels 8; Gaps 3;
Matches 27; Conservative 3; Mismatches 10;

```

QY 151 PTTTTPVTPTTMSIPTTTTPTMTVSTTTSVPTTTSIPTTTSVP 198
Db 2 PTTTTTTTTTTT---TTITKP-PITATTTKPPTTT---TTITKP 41

```

RESULT 2

US-09-060-767B-8
; Sequence 8, Application US/09060767B
; Patent No. 6720152
; GENERAL INFORMATION:
; APPLICANT: Weil, Gary
; APPLICANT: Chandrashekar, Ramaswamy
; TITLE OF INVENTION: Diagnosis of Histoplasmosis Using Antigens Specific for
; FILE OF INVENTION: H. capsulatum
; FILE REFERENCE: BJCH 9986
; CURRENT APPLICATION NUMBER: US/09/060,767B
; CURRENT FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/043,332
; PRIOR FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 41
; TYPE: PRT

```
; ORGANISM: Histoplasma Capsulatum
US-09-060-767B-8
Query Match          4.9%; Score 86.5; DB 4; Length 41;
Best Local Similarity 35.8%; Pred. NO. 2.1;
Matches 19; Conservative 8; Mismatches 11; Indels 15; Gaps 2;

QY 151 PTTTTPPTTMSIPTTTTPTTMTVSTTTSVPTTTSIPTTTSVPVTTTV 203
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 PTTTPTTTTPT-----PTPSI-----IPITPIVPANKTIIVLTITII 39

RESULT 3
US-09-043-731-23
; Sequence 23, Application US/09043731A
; Patent No. 6344203
; GENERAL INFORMATION:
; APPLICANT: The Austin Research Institute
; TITLE OF INVENTION: Mimicking Peptides in Cancer Therapy
; FILE REFERENCE: CALA-200
; CURRENT APPLICATION NUMBER: US/09/043,731A
; CURRENT FILING DATE: 1998-06-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: single
; OTHER INFORMATION: stranded linear peptide
US-09-043-731-23

Query Match          4.0%; Score 70.5; DB 3; Length 29;
Best Local Similarity 57.1%; Pred. NO. 29;
Matches 16; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 168 PTTTTPPTTMTVSTTTSVPTTTSIPTTT 195
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3 PTTTPISTT-TWVPTPTGTGTQTPPTT 29

RESULT 4
US-08-169-948B-24
; Sequence 24, Application US/08169948B
; Patent No. 5861271
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen
; APPLICANT: Collier, Katherine
; APPLICANT: Larenas, Edmund
; TITLE OF INVENTION: No. 5861271el Cellulase Enzymes and Systems
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/169,948B
; FILING DATE: DEC 17 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Horn, Margaret A.

QY 151 PTTTTPPTTMSIPTTTTPTTMTVSTTTSVPTTTSIPTTTSVPVTTTS 190
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 PGATTITSTRPP--SGPTTITRAT---STSSSTPTTSS 34

Query Match          3.8%; Score 67; DB 2; Length 34;
Best Local Similarity 47.5%; Pred. NO. 68;
Matches 19; Conservative 5; Mismatches 10; Indels 6; Gaps 2;

QY 151 PTTTTPPTTMSIPTTTTPTTMTVSTTTSVPTTTSIPTTTSVPVTTTS 190
      ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 PGATTITSTRPP--SGPTTITRAT---STSSSTPTTSS 34

RESULT 5
US-08-448-873-24
; Sequence 24, Application US/08448873
; Patent No. 5874276
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen
; APPLICANT: Collier, Katherine A.
; APPLICANT: Larenas, Edmund
; TITLE OF INVENTION: No. 5874276el Cellulase Enzymes and Systems
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 180 Kimball Way
; CITY: South San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/448,873
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/169,948
; FILING DATE: 17-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stone, Christopher L.
; REGISTRATION NUMBER: 35,696
; REFERENCE/DOCKET NUMBER: GC226D14
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 742-7555
; TELEFAX: (415) 742-7217
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-169-948B-24

Query Match          3.8%; Score 67; DB 2; Length 34;
Best Local Similarity 47.5%; Pred. NO. 68;
Matches 19; Conservative 5; Mismatches 10; Indels 6; Gaps 2;
```

QY 151 PTTTTPPTTTTMSIPTTTTPTMTVTSTTSVPTTTS 190
Db 1 PGATTITSTRPP--SGPTTITRAT-----STSSSTPTTSS 34

RESULT 6
US-08-382-452D-24
; Sequence 24, Application US/08382452D
; Patent No. 6268196
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Ward, Michael
; APPLICANT: Collier, Katherine D.
; APPLICANT: Larenas, Edmund A.
; TITLE OF INVENTION: NOVEL CELLULOSE ENZYMES AND SYSTEMS
; TITLE OF INVENTION: FOR THEIR EXPRESSION
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International
; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/382,452D
; FILING DATE: February 1, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Christopher L. Stone
; REGISTRATION NUMBER: 36,696
; REFERENCE/DOCKET NUMBER: GC226-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 742-7217
; TELEFAX: (415) 742-7555
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-382-452D-24

Query Match 3.8%; Score 67; DB 3; Length 34;
Best Local Similarity 47.5%; Pred. No. 68;
Matches 19; Conservative 5; Mismatches 10; Indels 6; Gaps 2;

QY 151 PTTTTPPTTTTMSIPTTTTPTMTVTSTTSVPTTTS 190
Db 1 PGATTITSTRPP--SGPTTITRAT-----STSSSTPTTSS 34

RESULT 7
US-08-507-362A-12
; Sequence 12, Application US/08507362A
; Patent No. 6562340
; GENERAL INFORMATION:
; APPLICANT: Bedford, Michael
; APPLICANT: Morgan, Andrew
; APPLICANT: Fowler, Timothy
; APPLICANT: Ward, Michael
; APPLICANT: Clarkson, Kathleen
; APPLICANT: Collier, Katherine
; APPLICANT: Larenas, Edmund
; TITLE OF INVENTION: An Enzyme Feed Additive and Animal Feed Including It
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genencor International

; STREET: 925 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/507,362A
; FILING DATE: 27-Oct-1995
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Castaneda, Janet
; REGISTRATION NUMBER: 33,228
; REFERENCE/DOCKET NUMBER: GC226-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 846-4072
; TELEFAX: (650) 845-6504
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 34 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-08-507-362A-12

Query Match 3.8%; Score 67; DB 4; Length 34;
Best Local Similarity 47.5%; Pred. No. 68;
Matches 19; Conservative 5; Mismatches 10; Indels 6; Gaps 2;

QY 151 PTTTTPPTTTTMSIPTTTTPTMTVTSTTSVPTTTS 190
Db 1 PGATTITSTRPP--SGPTTITRAT-----STSSSTPTTSS 34

RESULT 8
US-09-916-494A-24
; Sequence 24, Application US/09916494A
; Patent No. 6620605
; GENERAL INFORMATION:
; APPLICANT: Fowler, Timothy
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Ward, Michael
; APPLICANT: Collier, Katherine D.
; APPLICANT: Larenas, Edmund
; TITLE OF INVENTION: Method and Compositions for Treating
; TITLE OF INVENTION: Cellulose Containing Fabrics Using Truncated Cellulase
; TITLE OF INVENTION: Enzyme Compositions
; FILE REFERENCE: GC226-C4
; CURRENT APPLICATION NUMBER: US/09/916,494A
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 08/382,452
; PRIOR FILING DATE: 1995-02-01
; PRIOR APPLICATION NUMBER: US 08/169,948
; PRIOR FILING DATE: 1993-12-17
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 34
; TYPE: PRT
; ORGANISM: Trichoderma longibrachiatum
US-09-916-494A-24

Query Match 3.8%; Score 67; DB 4; Length 34;
Best Local Similarity 47.5%; Pred. No. 68;
Matches 19; Conservative 5; Mismatches 10; Indels 6; Gaps 2;

QY 151 PTTTTPPTTTTMSIPTTTTPTMTVTSTTSVPTTTS 190
Db 1 PGATTITSTRPP--SGPTTITRAT-----STSSSTPTTSS 34

APPLICANT: Weil, Gary
APPLICANT: Chandrasekar, Ramaswamy
TITLE OF INVENTION: Diagnosis of Histoplasmosis Using Antigens Specific for
FILE OF INVENTION: H. capsulatum
FILE REFERENCE: BJCH 9986
CURRENT APPLICATION NUMBER: US/09/060,767B
CURRENT FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/043,332
PRIOR FILING DATE: 1997-04-15
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 25
TYPE: PRT
ORGANISM: Caldocellum saccharolyticum
US-09-060-767B-4

Query Match 3.5%; Score 60.5; DB 4; Length 25;
Best Local Similarity 52.0%; Pred. No. 1.6e+02;
Matches 13; Conservative 1; Mismatches 10; Indels 1; Gaps 1;

QY 151 PTTTTPPTTPTTMSIPTTTTPT 175
DB 1 PTTTTPPTTPTTPTTPTTATPT 24

RESULT 13
US-08-530-569B-7
Sequence 7, Application US/08530569B
Patent No. 5939526
GENERAL INFORMATION:
APPLICANT: Gaugler, Beatrice
APPLICANT: van den Bynde, Benoit
APPLICANT: Schrier, Peter
APPLICANT: Brouwenstijn, Nathalie
APPLICANT: Boon-Fallieur, Thierry
TITLE OF INVENTION: Isolated RAGE-1 Derived Peptides Which
TITLE OF INVENTION: Complex with HLA-B7 Molecules and Uses Thereof
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: Federal Reserve Plaza, 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02210

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/530,569B
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Gates, Edward R.
REGISTRATION NUMBER: 31,616
REFERENCE/DOCKET NUMBER: L0461/7002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-530-569B-7

Query Match 3.4%; Score 59; DB 2; Length 40;
Best Local Similarity 42.9%; Pred. No. 3.8e+02;
Matches 12; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

QY 186 PTTTSIPTTSPVPTTSTTVPMPPLP 213
DB 6 PLRRSSPSSNRIRNTSTNNQFVPTMPLP 33

RESULT 14
US-08-597-495B-28
Sequence 28, Application US/08597495B
Patent No. 5712369
GENERAL INFORMATION:
APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerd;
APPLICANT: Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
APPLICANT: Catimel, B.; Ji, Hong; Burgess, Anthony W.;
APPLICANT: Heath, Joan K.; White, Sara J.; Johnstone, Cameron
TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
TITLE OF INVENTION: Associated Nucleic Acid Molecules, Protein And Peptides
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/597,495B
FILING DATE: 02-Feb-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/511,876
FILING DATE: 04-Aug-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5712369man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5316.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-597-495B-28

Query Match 3.4%; Score 59; DB 1; Length 41;
Best Local Similarity 29.4%; Pred. No. 3.9e+02;
Matches 10; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 74 DTRYKLLGLSRDVSILTIENTAVSDSGVYCCRV 107
DB 2 ENRVVSNDAEKSNASITIDQKTMDNGTYECSV 35

RESULT 15
US-09-068-051A-28
Sequence 28, Application US/09068051A
Patent No. 6291235
GENERAL INFORMATION:
APPLICANT: Old, Lloyd J.; Welt, Sydney; Ritter, Gerd;
Simpson, Richard J.; Nice, Edouard; Moritz, R. L.;
Catimel, B.; Ji, Hong; Burgess, Anthony W.;
Heath, Joan K.; White, Sara J.; Johnstone, Cameron
TITLE OF INVENTION: Colon Cell And Colon Cancer Cell
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fulbright & Jaworski LLP
STREET: 666 Fifth Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/068,051A
FILING DATE: 10-Dec-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/597,495
FILING DATE: 02-Feb-1996
APPLICATION NUMBER: 08/511,876
FILING DATE: 04-Aug-1995
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 6291235man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5316.2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 318-3168
TELEFAX: (212) 752-5958
INFORMATION FOR SEQ ID NO: 28
SEQUENCE CHARACTERISTICS:
LENGTH: 41 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 28
US-09-068-051A-28

Query Match 3.4%; Score 59; DB 3; Length 41;
Best Local Similarity 29.4%; Pred. No. 3.9e+02;
Matches 10; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 74 DTRYKLLGDLRRDVLNTENTAVSDSGVYCCRV 107
: : : : : : : : : : : : : : : :
Db 2 ENRVVSNDAEKNASITIDQKTMDDNGTCECV 35

Search completed: June 29, 2005, 09:49:08
Job time : 61.3694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:03:05 ; Search time 228.354 Seconds
(without alignments)
565.693 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLIHLADSVAG.....PFKKEVQQLRPHKSCIHQRE 334

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 938430

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	5.8	18	AAO26686	Human KIM
2	99	5.7	18	AAO26687	Human KIM
3	99	5.7	18	AAO26673	Monoclonal
4	98	5.6	18	AAO26681	Human KIM
5	97	5.5	18	AAO26682	Human KIM
6	96.5	5.5	49	AAW59911	Amino aci
7	96	5.5	18	AAO26683	Human KIM
8	95	5.4	18	AAO26685	Human KIM
9	92	5.3	18	AAO26684	Human KIM
10	90	5.1	18	AAO26688	Human KIM
11	86.5	4.9	39	ABU53167	Human tes
12	79.5	4.5	34	ABU53169	Human tes
13	73	4.2	45	ABP96634	Raw-starc
14	71	4.1	38	AAW72130	Molecular
15	70.5	4.0	29	AAW21983	Human MOC
16	70.5	4.0	29	AAW22257	Muc pep 9
17	70.5	4.0	43	ABB41070	Peptide #
18	70.5	4.0	43	AAW34846	Peptide #
19	70.5	4.0	43	AAW74730	Human bron
20	70.5	4.0	43	AAW61928	Human bron
21	70.5	4.0	43	ABG56513	Human liv
22	70.5	4.0	43	ABG44533	Human tes
23	67	3.8	38	ADG14245	Cellulose
24	67	3.8	46	AAW88522	Amphotrop
25	65	3.7	48	AAO19219	Human sec

26	65	3.7	48	5	AAO19218	Human sec
27	64.5	3.7	39	3	AAG10258	Arabidops
28	63	3.6	32	7	ADM08777	Canine im
29	63	3.6	32	7	ADM08560	Canine im
30	62	3.5	34	4	AAW17640	Peptide #
31	62	3.5	34	4	ABB36660	Peptide #
32	62	3.5	34	4	AAW30158	Peptide #
33	62	3.5	34	4	ABB31447	Peptide #
34	62	3.5	34	4	ABB21997	Protein #
35	62	3.5	34	4	AAW69820	Human bon
36	62	3.5	34	4	AAW57424	Human bra
37	62	3.5	34	4	ABG51510	Human liv
38	62	3.5	34	4	AAW05299	Peptide #
39	62	3.5	34	5	ABG39446	Human pep
40	61.5	3.5	26	2	AAW72699	Human muc
41	61	3.5	34	7	ADG14246	Cellulose
42	60	3.4	39	4	ABU53259	Human tes
43	60	3.4	50	4	AAW20331	Peptide #
44	60	3.4	50	4	ABB40865	Peptide #
45	60	3.4	50	4	AAW34632	Peptide #

ALIGNMENTS

RESULT 1
AAO26686
ID AAO26686 standard; peptide; 18 AA.
XX
AC AAO26686;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 14.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
(BIOJ) BIOGEN INC.
(GEHO) GEN HOSPITAL CORP.
XX
Bailly V, Bonventre J;
XX
WPI; 2003-156845/15.
XX
DR New antibody, antibody derivative or antigen-binding polypeptide that
XX inhibits proteolytic release of a soluble kidney injury molecule-1
XX polypeptide, useful for treating or preventing renal disease or injury,
XX e.g. renal cancer.
XX
PS Disclosure; Fig 1A; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
XX -binding polypeptide that inhibits proteolytic release of a soluble
XX kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
XX The antibody, antibody derivative or antigen-binding polypeptide is
XX useful for treating or preventing renal disease or injury, e.g. renal
XX cancer. The antibody is also useful for inhibiting shedding of the KIM-1
XX polypeptide. This sequence represents a human KIM-1 mucin domain related
XX peptide of the invention

```

SQ Sequence 18 AA;
Query Match
Best Local Similarity 5.8%; Score 101; DB 6; Length 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 TDGNDTVTSSDGLWNN 272
DB 1 TDGNDTVTSSDGLWNN 18

RESULT 2
AAO26687
ID AAO26687 standard; peptide; 18 AA.
XX
AC AAO26687;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 15.
XX
DE Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
XX
PR 04-JUN-2001; 2001US-0295907P.
XX
PA (BIOJ ) BIOGEN INC.
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Bailly V, Bonventre J;
XX
DR WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Claim 4; Page 25; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a binding epitope of the invention
XX
SQ Sequence 18 AA;
Query Match 5.7%; Score 99; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 SSDGLWNNQTQLFLEHS 281
DB 1 SSDGLWNNQTQLFLEHS 18

RESULT 4
AAO26681
ID AAO26681 standard; peptide; 18 AA.
XX
AC AAO26681;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 9.
XX
DE Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX

```


RESULT 7
 AAO26683
 ID AAO26683 standard; peptide; 18 AA.
 XX
 AC AAO26683;
 XX
 DT 20-MAR-2003 (first entry)
 XX
 DE Human KIM-1 mucin domain related peptide, SEQ ID No 11.
 XX
 KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KW renal disease; injury; renal cancer.
 XX
 OS Synthetic.
 XX
 PN WO200298920-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-US017402.
 XX
 PR 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Bailly V, Bonventre J;
 XX
 DR WPI; 2003-156845/15.
 XX
 PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-US017402.
 XX
 PR 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Bailly V, Bonventre J;
 XX
 DR WPI; 2003-156845/15.
 XX
 PT New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Disclosure; Fig 1A; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX
 SQ Sequence 18 AA;
 XX
 Query Match 5.5%; Score 96; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.2;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 228 QPAETHPTTLOGAIRREP 245
 DB 1 QPAETHPTTLOGAIRREP 18
 XX
 RESULT 8
 AAO26685
 ID AAO26685 standard; peptide; 18 AA.
 XX
 AC AAO26685;
 XX
 DT 20-MAR-2003 (first entry)
 XX
 DE Human KIM-1 mucin domain related peptide, SEQ ID No 13.
 XX
 KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KW renal disease; injury; renal cancer.

KW renal disease; injury; renal cancer.
 XX
 OS Synthetic.
 XX
 PN WO200298920-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-US017402.
 XX
 PR 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (GEO) GEN HOSPITAL CORP.
 XX
 PI Bailly V, Bonventre J;
 XX
 DR WPI; 2003-156845/15.
 XX
 PT New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Disclosure; Fig 1A; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX
 SQ Sequence 18 AA;
 XX
 Query Match 5.4%; Score 95; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 3.8;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 246 TSSPLYSYTTDGNPTVTE 263
 DB 1 TSSPLYSYTTDGNPTVTE 18
 XX
 RESULT 9
 AAO26684
 ID AAO26684 standard; peptide; 18 AA.
 XX
 AC AAO26684;
 XX
 DT 20-MAR-2003 (first entry)
 XX
 DE Human KIM-1 mucin domain related peptide, SEQ ID No 12.
 XX
 KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KW renal disease; injury; renal cancer.
 XX
 OS Synthetic.
 XX
 PN WO200298920-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-US017402.
 XX
 PR 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX
 PA (BIOJ) BIOGEN INC.

PA (GEHO) GEN HOSPITAL CORP.
 XX Bailly V, Bonventre J;
 XX WPI; 2003-156845/15.
 XX New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX Disclosure; Fig 1A; 42pp; English.
 XX The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX Sequence 18 AA;
 SQ

Query Match 5.3%; Score 92; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 6.2;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 237 LQAIRREPTSPLYSYT 254
 DB 1 LQAIRREPTSPLYSYT 18

RESULT 10
 AAO26688
 ID AAO26688 standard; peptide; 18 AA.
 XX AAO26688;
 XX 20-MAR-2003 (first entry)
 XX Human KIM-1 mucin domain related peptide, SEQ ID No 16.
 DE Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KW renal disease; injury; renal cancer.
 XX Synthetic.
 OS WO200298920-A1.
 PN 12-DEC-2002.
 PD 31-MAY-2002; 2002WO-US017402.
 PF 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX (BIOJ) BIOGEN INC.
 PA (GEHO) GEN HOSPITAL CORP.
 XX Bailly V, Bonventre J;
 XX WPI; 2003-156845/15.
 XX New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX Disclosure; Fig 1A; 42pp; English.
 XX The invention relates to a novel antibody, antibody derivative or antigen

CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX Sequence 18 AA;
 SQ

Query Match 5.1%; Score 90; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 8.6;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 273 QTQLFLEHSLLTANTTKG 290
 DB 1 QTQLFLEHSLLTANTTKG 18

RESULT 11
 ABUS3167
 ID ABUS3167 standard; protein; 39 AA.
 XX ABUS3167;
 AC ABUS3167;
 XX 14-APR-2003 (first entry)
 XX Human testes-derived DKFZphtes3_2a11 homologue #27.
 DE Human; gene therapy; vaccine; disease treatment; detection.
 KW Homo sapiens.
 OS WO200112659-A2.
 PN 22-FEB-2001.
 PD 18-AUG-2000; 2000WO-IB001496.
 PF 18-AUG-1999; 99US-0149499P.
 PR 28-SEP-1999; 99US-0156503P.
 XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
 PA Wiemann S;
 PI WPI; 2001-327840/34.
 DR Nucleic acids having the sequences of clones isolated from libraries of
 XX different human tissues, useful in recombinant DNA methodologies.
 XX Example III; Page 776; 1095pp; English.
 XX This invention describes novel polynucleotides and polypeptides isolated
 CC from human cDNA libraries which can be used for gene therapy or in
 CC vaccines. The polynucleotides of the invention and antibodies encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The products of the
 CC invention may also be used to identify modulators of expression and
 CC activity and to down regulate expression and activity. The antibodies of
 CC the invention may also be used as diagnostic agents for detecting the
 CC presence of polypeptides in samples. This sequence represents a homologue
 CC of a polypeptide described in the disclosure of the invention
 XX Sequence 39 AA;
 SQ

Query Match 4.9%; Score 86.5; DB 4; Length 39;
 Best Local Similarity 55.0%; Pred. No. 38;
 Matches 22; Conservative 5; Mismatches 10; Indels 3; Gaps 2;

QY 145 RSTST-TVPTTTTTPVTMTSIPITTTTTPVTMTVSTTT 183
 DB 1 RSTSTTTTITLPTT--PTSTTTTTPVTSTVLSTT 38


```

PT  Cancer vaccine containing MUC1 peptide mimic - used in the treatment of
XX  adenocarcinoma.
XX
XX  Example 1; Page 27; 55pp; English.
XX
CC  Muc pep9 is a peptide derived from the variable number of tandem repeat
CC  (VNTR) region of human cancer mucin MUC2. Unlike other mucin VNTR-derived
CC  peptides (AAW21975-85), it is not capable of binding to anti-Gal
CC  alpha(1,3)Gal antibody. A novel cancer vaccine comprises a peptide
CC  (AAW21680-86) which mimics MUC1 or other cancer peptides and one or more
CC  pharmaceutically acceptable carrier or diluent, optionally in association
CC  with an appropriate carrier peptide or other therapeutic agent. Some
CC  mucin VNTR peptides (AAW21975-82) may also be used. The vaccine is used
CC  in the treatment of a patient suffering from, or with a predisposition
CC  to, adenocarcinoma (claimed)
XX
XX  Sequence 29 AA;
SQ
    Query Match          4.0%; Score 70.5; DB 2; Length 29;
    Best Local Similarity 57.1%; Pred. No. 3.8e+02;
    Matches 16; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

Qy   168 PTTTVPPTTMVSTTSVPTTSIPPTT 195
      |||| : || : || || || ||
Db   3 PTTTPISTT-TMVTPTPTGTQTPTTT 29

Search completed: June 29, 2005, 09:41:07
Job time : 230.354 secs

```

FT		/note= "factor Xa cleavage recognition sequence"	
XX	JP06277088-A.		
PN	XX		
DD	XX		
PD	04-OCT-1994.		
XX			
XX	30-MAR-1993;	93JP-00071559.	
PF	XX		
PR	30-MAR-1993;	93JP-00071559.	
XX	(TOXS) TOYO ENG CORP.		
PA	XX		
XX	WPI; 1994-353764/44.		
DR	N-PSDB; AAQ89357.		
DR			
XX			
PT	Prod'n. and purificn. of recombinant protein without abnormal folding - by		
PT	adding fused protein to transformed host organism and purifying fused		
PT	protein by affinity chromatography using cellulose@ carrier.		
XX			
PS	Claim 2; Fig 3; 14pp; Japanese.		
XX			
CC	Recombinant proteins can be produced by affinity purification by		
CC	utilising a molecular affinity tag. This sequence shows a molecular tag		
CC	encoded by AAQ89357, and contains an endoglucanase III linker domain		
CC	(EGIII LD) and a factor Xa cleavage recognition site. A nucleotide		
CC	encoding a recombinant protein of interest can be inserted into a MCS		
CC	(multiple cloning site), downstream of the cleavage site. The fused		
CC	protein is purified when the EGIII LD is linked to an EGIII cellulose		
CC	binding domain which binds to cellulose and affinity chromatography is		
CC	carried out. The recombinant protein can then be recovered without the		
CC	tag by cleavage using a specific protease, eg. factor Xa		
XX			
SQ	Sequence 38 AA;		
	Query Match	4.1%; Score 71; DB 2; Length 38;	
	Best Local Similarity	48.8%; Pred. No. 4.8e+02;	
	Matches 20; Conservative	5; Mismatches 10; Indels	6; Gaps 2;
Qy	151 PTTTTPVTTTVPTTMSIPTTTTVTPTTWSTTTVPPTTSI	191	
	:		
Dd	1 PGATTITSTRPP--SGPTTTTRAT---STSSSPPTSSI	35	

RESULT 15	
AAW21983	
ID	AAW21983 standard; peptide; 29 AA.
XX	
XX	AAW21983;
AC	
XX	
DT	30-OCT-1997 (first entry)
XX	
DE	Human MUC2 VNTR peptide Muc pep9.
XX	
KW	Cancer; vaccine; peptide mimic; mucin; MUC2; Gal alpha(1,3)Gal;
KW	immunotherapy; therapy; adenocarcinoma; VNTR;
KW	variable number of tandem repeat.
XX	
OS	Homo sapiens.
XX	
PN	WO9711715-A1.
XX	
PD	03-APR-1997.
XX	
PF	27-SEP-1996; 96WO-AU000617.
XX	
PR	27-SEP-1995; 95AU-00005680.
XX	
PA	(AUST-) AUSTIN RES INST.
XX	
PI	Sandrin MS, McKenzie IFC, Apostolopoulos V;
XX	
DR	WPI; 1997-212670/19.
XX	

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 30.1847 Seconds
(without alignments)
1064.659 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLILHLADSVAG.....FPKKEVQQLRPHKSCIHORE 334
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1309	74.8	451	S71754	cellular hepatitis
2	260	14.9	770	T22808	hypothetical prote
3	246.5	14.1	662	A45155	mucin FIM-C.1 - Af
4	244	14.0	400	A28172	spasmolysin precu
5	242.5	13.9	851	T22696	hypothetical prote
6	234	13.4	371	S20075	promastigote surfa
7	229	13.1	216	I51920	mucin - rhesus mac
8	227.5	13.0	1832	T31113	mucin-like glycopor
9	221.5	12.7	3020	A43932	mucin 2 precursor,
10	219.5	12.6	327	S20074	promastigote surfa
11	216.5	12.4	294	A37232	mucin, tracheal (A
12	215	12.3	232	A60095	larval glue protei
13	214.5	12.3	235	PC2022	mucin like protein
14	213	12.2	167	A33532	mucin SMUC-40 - hu
15	212.5	12.1	307	GSF3	salivary glue prot
16	206	11.8	1161	S57180	probable membrane
17	205	11.7	447	A39321	mucin - rat (fragm
18	202.5	11.6	825	T29634	hypothetical prote
19	201.5	11.5	379	S50125	larval glue protei
20	201	11.5	660	JW0067	chitinase (EC 3.2
21	200.5	11.5	1118	A48292	mucin, tracheobron
22	198	11.3	1513	A54895	mucin 2, intestina
23	195	11.1	345	E88103	protein W10G11.5 l
24	194	11.1	693	T19551	mucin-like protein
25	192.5	11.0	708	T19474	hypothetical prote
26	192.5	11.0	796	T21460	hypothetical prote
27	191	10.9	798	T34248	hypothetical prote
28	190.5	10.9	592	T34446	hypothetical prote
29	189.5	10.8	474	S15921	protein TPX-VT3 -

30	188.5	10.8	977	2	T16232	hypothetical prote
31	187.5	10.7	797	1	VBEX1	glycoprotein x pre
32	187.5	10.7	1235	2	T13710	protein-tyrosine k
33	186.5	10.7	292	2	S24169	mucin - rat
34	185.5	10.6	263	2	S01360	salivary glue prot
35	185	10.6	605	2	S48940	hypothetical prote
36	184	10.5	795	2	T20609	hypothetical prote
37	182.5	10.4	790	2	T34293	hypothetical prote
38	181.5	10.4	279	2	S53363	mucin SAC (clone J
39	181.5	10.4	602	2	AD2067	hypothetical prote
40	180	10.3	1777	2	T34369	hypothetical prote
41	179.5	10.3	921	2	AB0332	conserved hypotet
42	178	10.2	477	2	S53362	mucin SAC (clone J
43	177	10.1	560	2	T32661	hypothetical prote
44	175.5	10.0	217	2	S01358	salivary glue prot
45	175	10.0	866	2	T45462	membrane glycoprot

ALIGNMENTS

RESULT 1

S71754
cellular hepatitis A receptor HAVcr-1 precursor - green monkey
N:Alternate names: surface glycoprotein
C:Species: Cercopithecus aethiops (green monkey, grivet)
C:Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C:Accession: S71754

R:Kaplan, G.; Totsuka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstone, S.M.
EMBO J. 15, 4282-4296, 1996
A:Title: Identification of a surface glycoprotein on African green monkey kidney cells as
A:Reference number: S71754; MUID:97015129; PMID:8861957
A:Accession: S71754
A:Molecule type: mRNA

A:Residues: 1-451 <KAP>
A:Cross-references: UNIPROT:Q95144; EMBL:X98252; NID:g1526573; PID:e247449; PID:g1526574
A:Experimental source: kidney
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match 74.8%; Score 1309; DB 2; Length 451;
Best Local Similarity 63.3%; Pred. No. 8.8e-74;
Matches 255; Conservative 36; Mismatches 32; Indels 80; Gaps 3;

Qy	1	MHPQVILSLILHLADSVAGSVKVG3EAGPSVTLPCYSGAVTSMCNRGSCSLFTCONG	60
Db	1	MHPQVILSLILHLADSVAGSVKVG3EAGPSVTLPCYSGAVTSMCNRGSCSLFTCONG	60
Qy	61	IWTNGTHVYRKOTRYKLLGDLSDRRDVSLLTENTAVSDSGVYCCRVHRGFNDMKITV	120
Db	61	IWTNGTHVYRKOTRYKLLGDLSDRRDVSLLTENTAVSDSGVYCCRVHRGFNDMKITV	120
Qy	121	SLEIVPPKV-----TTTPIVTTVP	139
Db	121	SLKIGPRVTPIVRTVSTVPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPT	180
Qy	140	TVTVTSTVPTTTTTP-----TTTVPMTMSIPTTTTVPMTVSTT	182
Db	181	TTTTVPTTTTLPTTTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPT	240
Qy	183	T-----SVPTTTSIPTTTSVPTTTTVPVTPVPMPLPRQNHFPV	220
Db	241	TTLPPTTLPTMTLPTTTTLPTMTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPT	300
Qy	221	ATSPSSPOPAETHPTLOGAIRRBTSSPLYSYTTGNDTDTSSDGLNNTQTLFLXH	280
Db	301	ATSPSSPOPAETHPTVLLGATRTQTPSSPLSYTTGSDTDTSSDGLNNTQTLSPH	360
Qy	281	SLLTANTTKGIYAGVCISVLVLLGVLIAKKYFKKEVQOL	323
Db	361	SPQWNTTEGIYAGVCISVLVLLGVLIAKKYFKKEIQOL	403

```
RESULT 2
T22808
hypothetical protein F56H9.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T22808
R:Burton, J.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19618
A:Accession: T22808
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-770 <WIL>
A:Cross-references: UNIPROT:Q20908; EMBL:Z74473; PIDN:CAA98949.1; GSPDB:GN000023; CESP:F56H9
A:Experimental source: clone F56H9
C:Genetics:
A:Gene: CESP:F56H9.1
A:Map position: 5
A:Introns: 235/1; 262/2; 320/1; 367/2; 510/3; 654/1; 681/2
Query Match 14.9%; Score 260; DB 2; Length 770;
Best Local Similarity 40.9%; Pred. No. 7.6e-09;
Matches 79; Conservative 18; Mismatches 66; Indels 30; Gaps 8;
QY 119 TVSLIIVPKVTTPIVTV--PTVTVRSTTVPTTVPTTV---PTTMSIPTTTVP 174
DB 87 TTTTIVPTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTTSTTTTTP 146
QY 175 TMTVSTTTSV--PTTTSIPTTSPVPTT---VSTFVPPMPLPRQNHPEVATSPSSQP 229
DB 147 PTTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTT 197
QY 230 AETHPTLQGAIRREPTSPVLSYTTDGDVTE-----SSD-----GLMNNQQLFLE 279
DB 198 PTTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTTSTTTTTPPTT 257
QY 280 --HSLLTANTTKG 290
DB 258 VTYSLNLANIIPG 270
RESULT 3
A45155
mucin FIM-C.1 - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C>Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 09-Jul-2004
C:Accession: A45155
R:Hauger, F.; Hoffmann, W.
J. Biol. Chem. 267, 24620-24624, 1992
A:Title: P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1)
A:Reference number: A45155; MUID:93077556; PMID:1447205
A:Accession: A45155
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-662 <HAU>
A:Cross-references: UNIPROT:Q05049; GB:I02115; NID:g214147; PIDN:AAA74725.1; PID:g951460
F:162-202/Domain: trefoil homology <TRF1>
F:307-347/Domain: trefoil homology <TRF2>
F:354-394/Domain: trefoil homology <TRF3>
F:526-566/Domain: trefoil homology <TRF4>
F:573-613/Domain: trefoil homology <TRF5>
F:621-661/Domain: trefoil homology <TRF6>
Query Match 14.1%; Score 246.5; DB 2; Length 662;
Best Local Similarity 30.4%; Pred. No. 4.4e-08;
Matches 77; Conservative 25; Mismatches 90; Indels 61; Gaps 9;
QY 30 PSVTLFCHYGAVTSMCWNRGCSLFTQCGIWTNGTHVT-----YRKDRFY 77
DB 311 PSKREDCGYSGITGSCQRTKGCC----FDSSIPQTKWCFTYLSQVADCKVPSQRVDCGF 366
QY 78 K-LLGLDSRR-----DVSLSLTIENTAVSDSGVCCRVEHGWFNMDMKITVSLIIVPKVTT 131
```

```
Db 367 RCITADQCQRKCCFDSSI-----SGIKWC-----FVSTSQVAATKTTTPTTTT 411
QY 132 TPVIVTV--PTVTVRSTTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 189
DB 412 TPTTTTTPKATTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 471
QY 190 SIPTTTSVPVTTTSTFVPPMPLPRQNHPEVATSPSSQPAAETHPTTLQGAIRREPTSSP 249
DB 472 TPTTPTTTPKATTTPTT-----TTTTTPTTTTKAAT-----TPTPTTTT 509
QY 250 LYSVTTDGDNDVTV 262
DB 510 TTTTTPKATTTT 522
RESULT 4
A28172
spasmolysin precursor - African clawed frog
N:Alternate names: prospasmolysin
C:Species: Xenopus laevis (African clawed frog)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A28172
R:Hoffmann, W.
J. Biol. Chem. 263, 7686-7690, 1988
A:Title: A new repetitive protein from Xenopus laevis skin highly homologous to pancreatrin
A:Reference number: A28172; MUID:88227968; PMID:3372504
A:Accession: A28172
A:Molecule type: mRNA
A:Residues: 1-400 <HOF>
A:Cross-references: UNIPROT:P10667; GB:M19971; NID:g214791; PIDN:AAA49960.1; PID:g214792
C:Superfamily: spasmolysin; trefoil homology
C:Keywords: duplication; glycoprotein; skin; tandem repeat
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-400/Product: spasmolysin #status predicted <MAT>
F:21-69/Product: spasmolysin I #status predicted <SP1>
F:23-63/Domain: trefoil homology <TRF1>
F:72-348/Product: spasmolysin-glycopeptide #status predicted <SPG>
F:74-114/Domain: trefoil homology <TRF2>
F:115-309/Region: PEST sequence
F:123-131,231-239/Region: 9-residue repeats (D/E-T-T-T-A-S-T-T-A)
F:132-230,240-248/Region: 9-residue repeats (E-T-T-T-V-P-T-T-P)
F:249-252,276-299/Region: 4-residue repeats (E-T-T-T-T)
F:305-342/Domain: trefoil homology <TRF3>
F:351-400/Product: spasmolysin II #status predicted <SP2>
F:353-393/Domain: trefoil homology <TRF4>
F:23-49,33-48,43-60,74-100,84-99,94-111,312-327,322-339,353-379,363-378,373-390/Disulfide
F:63/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 14.0%; Score 244; DB 1; Length 400;
Best Local Similarity 31.9%; Pred. No. 3.8e-08;
Matches 83; Conservative 24; Mismatches 107; Indels 46; Gaps 12;
QY 16 DSVAGSV-KVGGEAGPSVTLFCHYGAVTSMCWNRGCSLFTQCGIWTNGTHV--TYR 72
DB 63 NATAGPIKLCSCGPTTKRIDGPRITEKQICILRG-CFDDSSISGVKWCYARTVITTPA 121
QY 73 KDTRYKLLGDLSSRDVSLTIENTAVSDSGVCCRVEHGWFNMDMKITVSLIIVPKVT-- 130
DB 122 PDT-----TTASTTAAETTTVPPT-----PETTVPTTPTTVP 155
QY 131 TPIVTVPTTVTRTSTTVPTTTPV-----TTVPTTMSIPTTTTPTTMTVSTTTSVP 186
DB 156 TPTTIVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 212
QY 187 TTTTSIPTTSPVTTTSTFVPPMPLPRQNHPEVATSPSSQPAAETHPTTLQGAIRREPT 246
DB 213 ETTTVPTT--PEITTVPT--TPETTTASTTAAETTTVPTTBE-TTTEPTTPTTDTTPTT 266
QY 247 SSPLSYTVD-GNDVTLESS 265
DB 267 LPPTPETTTTETTTTETT 286
```

[illegible]

mucin SMUC-40 - human (fragment)
C/Species: Homo sapiens (man)
C/Date: 29-Jan-1990 #sequence_revision 27-Jun-1994 #text_change 09-Jul-2004
C/Accession: A33532
R/Gum, J.R.; Byrd, J.C.; Hicks, J.W.; Toribara, N.W.; Lampert, D.T.A.; Kim, Y.S.
J. Biol. Chem. 264, 6480-6487, 1989
A/Title: Molecular cloning of human intestinal mucin cDNAs. Sequence analysis and evidence for a family of genes.
A/Reference number: A33532; MUID:89197956; PMID:2703501
A/Accession: A33532
A/Molecule type: mRNA
A/Residues: 1-167 <GUN>
A/Cross-references: UNIPROT:Q14883; GB:M22404; NID:g188865; PIDN:AAA36333.1; PID:g188866
C/Keywords: glycoprotein

Query March	12.2%	Score 213;	DB 2;	Length 167;
Best Local Similarity	41.0%;	Pred. No. 1.3e-06;		
Matches 57;	Conservative 11;	Mismatches 51;	Indels 20;	Gaps 6;
Qy	124	IVPKVTTTPTVTVPTVTVTSTVTP- -TTTTPVTTVPTTWSIPTTTTTPVTMTVST	181	
Db	26	LTPTPTTTTTVTPPTPTGTPTSTPTSTPTSTPTPTPTPTPTPTPTPTPTTTT-T	84	
Qy	182	TTSPPTTTSIPTTTSVPVTTTTSFVPPMLPRQNHEP- - - - -VATSSSPQPAETHPTT	236	
Db	85	PTPTPTGTQTPPT- -PTTT- -TTVPTPTPTGTQTPPTPTPTTTTPTPTPTGTQT	140	
Qy	237	LOGAIRREPTSSPLYSYTT	255	
Db	141	-----PTSTPTTNTT	151	

```

RESULT 15
GSFF3
salivary glue protein sgs-3 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 09-Jul-2004
C:Accession: A03329
R:Garfinkel, M.D.; Pruitt, R.E.; Meyerowitz, E.M.
J. Mol. Biol. 168, 765-789, 1983
A:Title: DNA sequences, gene regulation and modular protein evolution in the Drosophila
A:Reference number: A92904; MUID:83294545; PMID:6411930
A:Accession: A03329
A:Molecule type: DNA
A:Residues: 1-307 <GAR>
A:Cross-references: UNIPROT:P02840; GB:X01918; NID:g8581; PIDN:CAA25994.1; PID:g603989
C:Comment: This protein is produced by third-instar larvae.

```

	Query Match	12.1%	Score 212.5	DB 1	Length 307
	Best local Similarity	38.0%	Pred. No. 2.5e-06		
	Matches 62	Conservative 14	Mismatches 64	Indels 23	Gaps 8
Qy	119	TVSLLEIVPPKVTTPPIVTTVP	---VTVTRSTTVPTTT-----	TVPTTTPVPTMS	166
Db	80	TTQLPCVTP	--TTTKATTTKTTTATTTKATTTTKQTTLQLPCTTPTTTKQT	TTQ	137
Qy	167	IP-----TTTTVPT	---MTVSPTTTSVP	TTTSPTTSVPVTTTSTFVPPMPLRQNH	220
Db	138	LPCTTPPTTKPTTKPTTKPTTK	PTTK-PTTK-PTTKPTTKPTTKPTTK	-KPTT-KPTT	194
Qy	221	ATSPSSPOPETHPTTLQAIRREPT	SSPLYSYTTDGN	DTVE	263
Db	195	TKKPTTKPTTKPTTKPTTKPTTK	PTTKPTTKPTTKPTTKPTTKPTTK	PTTK	237

Search completed: June 29, 2005, 08:58:14
Job time : 32.1847 secs

A;Molecule type: DNA
A;Residues: 1-639 <BOE>

C;Genetics:
A;Gene: SGD:TOM71; TOM72
A;Cross-references: SGD:S0001159; MIPS:YHR117w
A;Map position: 8R
A;Genome: nuclear
C;Superfamily: mitochondrial outer membrane protein, 70K; tetraatricoptide repeat homol
C;Keywords: mitochondrial outer membrane; mitochondrion; transmembrane protein
F;86-119/Domain: tetraatricoptide repeat homology <TT01>
F;127-160/Domain: tetraatricoptide repeat homology <TT02>
F;161-194/Domain: tetraatricoptide repeat homology <TT03>
F;345-377/Domain: tetraatricoptide repeat homology #status atypical <TT04>
F;378-411/Domain: tetraatricoptide repeat homology <TT05>
F;412-445/Domain: tetraatricoptide repeat homology <TT06>
F;446-479/Domain: tetraatricoptide repeat homology <TT07>
F;480-513/Domain: tetraatricoptide repeat homology <TT08>
F;530-563/Domain: tetraatricoptide repeat homology <TT09>
F;564-597/Domain: tetraatricoptide repeat homology <TT10>

Query Match 22.7%; Score 62.5; DB 2; Length 639;
Best Local Similarity 40.5%; Pred. No. 13;
Matches 17; Conservative 6; Mismatches 4; Indels 15; Gaps 2;

Qy 2 LYSVTDTG---NDTVTESSDGLWNNQTQLFLEHSLTLTANT 39
||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 305 LYSATDEGYLVANDLLTKSTD-----MYHSLSLTSANT 335

RESULT 3
A90267
proteinase related protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: A90267
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
arrett, R.A.; Ragan, M.A.; SENSEN, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: A90267
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1269 <KUR>
A;Cross-references: UNIPROT:Q97Z06; GB:A8006641; NID:g13814328; PIDN:AAK41392.1; GSPDB:G
C;Genetics:
A;Gene: SS01141

Query Match 22.7%; Score 62.5; DB 2; Length 1269;
Best Local Similarity 29.2%; Pred. No. 29;
Matches 19; Conservative 10; Mismatches 19; Indels 17; Gaps 2;

Qy 4 SYTTDG-----NDTVTE-SSDGLWNNQTQLFLEHSLTLTNTKGIVAG 46
||| : ||| : ||| : ||| : ||| : ||| : ||| :
Db 1163 SFTHGTHTVINTQNPDPGGYNLTVTAIOSDGLSSNSLYLPENGLTNLNTKVNVISN 1222
.: .:

Qy 47 VCISV 51
: |

Db 1223 QLTVV 1227

RESULT 4
Tl8619
hypothetical protein AH6.10 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: Tl8619
R;Berks, M.
submitted to the EMBL Data Library, January 1995
A;Reference number: Z18998
A;Accession: Tl8619
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-329 <WIL>

Query Match 21.6%; Score 59.5; DB 1; Length 103;
Best Local Similarity 37.8%; Pred. No. 3.9;
Matches 17; Conservative 6; Mismatches 17; Indels 5; Gaps 2;

QY 3 YSYTTDGDNDVTESSDGLWNNQTQLFLEHSLLTANTTKGIYAGV 47
 ||||| : | : | : | : | : | : | : | : |
Dd 46 YSY-TDAN-----KSKGIVMNTLMEYLENPKKIPTGTMTIPAGI 85
 :|||:

RESULT 9

E86797
prophage p13 protein 07 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C>Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86797
R:Bolojin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich Genome Res. 11, 731-753, 2001
A>Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: E86797
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <STO>
A:Cross-references: UNIPROT_Q9CFT4; GB:AEO05176; PID:g12724367; PIDN:AAK05479.1; GSPPDB:GN
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: pi307

Query Match 21.6%; Score 59.5; DB 2; Length 520;
Best Local Similarity 28.3%; Pred. No. 24;
Matches 15; Conservative 10; Mismatches 21; Indels 7; Gaps 2;

QY 3 YSYTTDGDNDVTESSDGL-----WNNOQLFLEHSLLTANT--TKGIYAGVC 48
 ||||| : | : | : | : | : | : | : | : |
Dd 361 YATGLNCININTDFSSGMKYTRAVTNQNANNFIQSVMKNAPSVGTVVIC 413
 |||||:

RESULT 10

C82135
chitinase VC1952 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: C82135
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Winn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P., l., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: AB2035; MUID:20406833; PMID:10952301
A:Accession: C82135
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-846 <HEI>
A:Cross-references: UNIPROT_Q9KQP6; GB:AB004271; GB:AEO03852; NID:g9656490; PIDN:AAF9S10
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1952
A:Map position: 1

Query Match 21.5%; Score 59; DB 2; Length 846;
Best Local Similarity 30.9%; Pred. No. 49;
Matches 17; Conservative 6; Mismatches 18; Indels 14; Gaps 2;

QY 5 YTDDGDNDVTESSDGLWN-----NNOTQLFLEHSLLTANTTKGIYAGV 47
 ||||| : | : | : | : | : | : | : | : |
Dd 54 YWTQGNDPVTHS--GWDAWQILQCDCGGANNPFQVSITQSPFLNNAKIPQQSVVGL 106
 |||||:

RESULT 11

JQ1978
structural polyprotein - Venezuelan equine encephalitis virus (subtype II, strain Everglades) N; Contains: 6K protein; coat protein; membrane glycoprotein EI; membrane glycoprotein E2; C:Species: Venezuelan equine encephalitis virus

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: JQ1978
C;Superfamily: togavirus structural polyprotein
J. Gen. Virol. 74, 519-523, 1993
A;Title: Molecular evidence that epizootic Venezuelan equine encephalitis (VEE) I-AB virus is a novel M-3 muscarinic acetylcholine receptor is expressed in chick atrium and ventricle
A;Reference number: JQ1978; MUID:93187617; PMID:8445371
A;Accession: JQ1978
A;Molecule type: mRNA
A;Residues: 1-1254 <SNE>
A;Cross-references: UNIPROT:P36330; GB:L04598; NID:G290612; PIDN:AAA42984.1; PID:G290614
C;Superfamily: togavirus structural polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-274/Product: coat protein #status predicted <CTP>
F;275-333/Product: membrane glycoprotein E3 #status predicted <MG3>
F;334-756/Product: membrane glycoprotein E2 #status predicted <MG2>
F;701-721/Domain: transmembrane #status predicted <TM1>
F;757-812/Product: 6K protein #status predicted <KP6>
F;794-813/Domain: transmembrane #status predicted <TM2>
F;813-1254/Product: transmembrane glycoprotein E1 #status predicted <MG1>
F;1231-1248/Domain: transmembrane #status predicted <TM3>
F;47,285,545,651,946/Binding site: carbohydrate (Asn) #status predicted

Query Match 21.5%; Score 59; DB 1; Length 1254;
Best Local Similarity 45.5%; Pred. No. 77;
Matches 15; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 11 DTVTSSDGLWNNQ-----TQLFLEHSLTANT 39
: ||| ||| ||| ||| ||| : : : :
DB 757 ETTWESLDHLWNNQMFWTQLLIPLAALIVT 789

RESULT 12
T45053
Hypothetical protein Y3986B.aa [imported] - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T45053
R;Wilson, R.; Ainscough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; C. A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johnson, B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifkin, L.; Roopra, A.; Saunders, D.; Nature 368, 32-38, 1994
A;Authors: Shownkeen, R.; Sims, M.; Smalton, N.; Smith, A.; Smith, M.; Sonhammer, E.; S. tock, L.; Wilkinson-Sproat, J.; Wohldman, P.
A;Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A;Reference number: T45053; MUID:94150718; PMID:7906398
A;Accession: T45053
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-321 <WIL>
A;Cross-references: UNIPROT:O9NET3; EMBL:AL132896; NID:G6434440; PIDN:CAB60932.1; PID:G6434440
A;Experimental source: clone Y3986B
C;Genetics:
A;Map position: 3
A;Introns: 41/2; 95/2; 131/1; 170/3; 208/1; 236/2; 279/3
A;Note: Y3986B.aa

Query Match 21.1%; Score 58; DB 2; Length 321;
Best Local Similarity 27.8%; Pred. No. 21;
Matches 15; Conservative 13; Mismatches 14; Indels 12; Gaps 2;

QY 2 LYSYTTDGDNTVTSSD---GLWNNQTLFLEHSLTANTTKGIYAGVCISV 51
: ||| ||| ||| ||| ||| : : : :
DB 33 MYTYITDNTDSEYANRAGCGVW-----AKYGIFTANIVFLVGGLLIAM 78

RESULT 13
A55019
muscarinic acetylcholine receptor, M3 isoform - chicken
C;Species: Gallus gallus (chicken)
C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C;Accession: A55019
J. Gaddbut, A. P.; Galper, J. B.
J. Biol. Chem. 269, 25823-25829, 1994

C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C;Accession: JQ1978
C;Superfamily: togavirus structural polyprotein
J. Gen. Virol. 74, 519-523, 1993
A;Title: Molecular evidence that epizootic Venezuelan equine encephalitis (VEE) I-AB virus is a novel M-3 muscarinic acetylcholine receptor is expressed in chick atrium and ventricle
A;Reference number: JQ1978; MUID:93187617; PMID:8445371
A;Accession: JQ1978
A;Molecule type: mRNA
A;Residues: 1-1254 <SNE>
A;Cross-references: UNIPROT:P36330; GB:L04598; NID:G290612; PIDN:AAA42984.1; PID:G290614
C;Superfamily: togavirus structural polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-274/Product: coat protein #status predicted <CTP>
F;275-333/Product: membrane glycoprotein E3 #status predicted <MG3>
F;334-756/Product: membrane glycoprotein E2 #status predicted <MG2>
F;701-721/Domain: transmembrane #status predicted <TM1>
F;757-812/Product: 6K protein #status predicted <KP6>
F;794-813/Domain: transmembrane #status predicted <TM2>
F;813-1254/Product: transmembrane glycoprotein E1 #status predicted <MG1>
F;1231-1248/Domain: transmembrane #status predicted <TM3>
F;47,285,545,651,946/Binding site: carbohydrate (Asn) #status predicted

Query Match 21.1%; Score 58; DB 2; Length 639;
Best Local Similarity 40.0%; Pred. No. 47;
Matches 18; Conservative 4; Mismatches 15; Indels 8; Gaps 3;

QY 7 TDGNDTVTSSDGLWNNQTLFLEHSL-----LTANTTKGIYA 45
: ||| ||| ||| ||| ||| : : : :
DB 372 TDQGDQEHSSSDS-WNNNDAAASSENSASDEEDITAE-TRAIYS 414

RESULT 14
VHWVVE
structural polyprotein - Venezuelan equine encephalitis virus (strain TC-83)
N;Contains: 6K peptide; coat protein; membrane glycoprotein E1; membrane glycoprotein E2;
C;Species: Venezuelan equine encephalitis virus
C;Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004
C;Accession: A27871
R;Johnson, B.J.B.; Kinney, R.M.; Kost, C.L.; Trent, D.W.
J. Gen. Virol. 67, 1951-1960, 1986
A;Title: Molecular determinants of alphavirus neurovirulence: nucleotide and deduced protein sequences
A;Reference number: A27871; MUID:86306669; PMID:3755750
A;Accession: A27871
A;Molecule type: mRNA
A;Residues: 1-1254 <JOH>
A;Cross-references: UNIPROT:P05674; GB:X04368; NID:G58621; PIDN:CAA27883.1; PID:G58622
A;Note: the authors translated the codon UUC for residues 19 and 27 as Ile and ACC for Leu
C;Superfamily: togavirus structural polyprotein
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-275/Product: coat protein #status predicted <COP>
F;276-334/Product: membrane glycoprotein E3 #status predicted <GP3>
F;335-757/Product: membrane glycoprotein E2 #status predicted <GP2>
F;702-718/Domain: transmembrane #status predicted <TM1>
F;758-812/Product: 6K protein #status predicted <PIP>
F;774-790/Domain: transmembrane #status predicted <TM2>
F;792-808/Domain: transmembrane #status predicted <TM3>
F;813-1254/Product: membrane glycoprotein E1 #status predicted <GP1>
F;1231-1247/Domain: transmembrane #status predicted <TM4>
F;47,286,546,625,946/Binding site: carbohydrate (Asn) #status predicted

Query Match 21.1%; Score 58; DB 1; Length 1254;
Best Local Similarity 47.6%; Pred. No. 1e+02;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 11 DTVTSSDGLWNNQTLFLE 31
: ||| ||| ||| ||| : : : :
DB 758 ETTWESLDHLWNNQMFVIQ 778

RESULT 15
VHWVVT
structural polyprotein - Venezuelan equine encephalitis virus (strain TRD)
N;Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2;
C;Species: Venezuelan equine encephalitis virus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C;Accession: B31467; A47612
R;Kinney, R.M.; Johnson, B.J.B.; Welch, J.B.; Tsuchiya, K.R.; Trent, D.W.
Virology 170, 19-30, 1989
A;Title: The full-length nucleotide sequences of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus
A;Reference number: A31467; MUID:89243175; PMID:2524126
A;Accession: B31467
A;Molecule type: mRNA
A;Residues: 1-1254 <KIL>
A;Cross-references: UNIPROT:P09592; GB:J04332; NID:G323708; PIDN:AAB02519.1; PID:G323710
R;Kinney, R.M.; Johnson, B.J.B.; Brown, V.L.; Trent, D.W.

Virology 152, 400-413, 1986
A;Title: Nucleotide sequence of the 26 S mRNA of the virulent Trinidad donkey strain of
A;Reference number: A47612; MUID:86263392; PMID:3088830
A;Accession: A47612
A;Molecule type: mRNA
A;Residues: 1-542,'K',544-810,'P',812-1254 <K12>
A;Cross-references: GB:L01442
C;Superfamily: togavirus structural polypeptide
C;Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F;1-275/Product: coat protein #status predicted <CTP>
F;276-334/Product: membrane glycoprotein E3 #status predicted <MG3>
F;335-757/Product: membrane glycoprotein E2 #status predicted <MG2>
F;701-718/Domain: transmembrane #status predicted <TM1>
F;758-812/Product: sK protein #status predicted <KP6>
F;774-790/Domain: transmembrane #status predicted <TM2>
F;795-813/Domain: transmembrane #status predicted <TM3>
F;813-1254/Product: membrane glycoprotein E1 #status predicted <MG1>
F;1231-1248/Domain: transmembrane #status predicted <TM4>
F;47,286,546,652,946/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.1%; Score 58; DB 1; Length 1254;
Best Local Similarity 47.6%; Pred. No. 1e+02;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 11 DTVTSSDGLWNNQQLFLE 31
Db 758 ETTWESLDHLWNNNQMFMIQ 778

Search completed: June 29, 2005, 08:58:15
Job time : 5.69941 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:58:20 ; Search time 209.324 Seconds
(without alignments)
613.589 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLIHLADSVAG.....PFKKEVQQLRPHKSCIHORE 334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1749	100.0	334	17	US-10-655-506-7
2	1749	100.0	334	17	US-10-718-321-7
3	1684	96.3	359	14	US-10-188-012-17
4	1684	96.3	359	15	US-10-295-027-302
5	1684	96.3	359	15	US-10-188-832-64
6	1684	96.3	359	17	US-10-391-939A-4
7	1684	96.3	359	17	US-10-663-497-17
8	1684	96.3	359	17	US-10-718-321-8
9	1684	96.3	359	17	US-10-847-918-25
10	1679	96.0	359	14	US-10-188-012-19
11	1679	96.0	359	17	US-10-663-497-19
					Sequence 7, Appli
					Sequence 7, Appli
					Sequence 17, Appl
					Sequence 302, App
					Sequence 4, Appli
					Sequence 17, Appl
					Sequence 8, Appli
					Sequence 25, Appl
					Sequence 19, Appl

12 1671 95.5 359 14 US-10-188-012-23 Sequence 23, Appl
13 1671 95.5 359 17 US-10-663-497-23 Sequence 23, Appl
14 1666 95.3 365 14 US-10-188-012-21 Sequence 21, Appl
15 1666 95.3 365 17 US-10-663-497-21 Sequence 21, Appl
16 1655.5 94.7 364 14 US-10-188-012-25 Sequence 25, Appl
17 1655.5 94.7 364 17 US-10-663-497-25 Sequence 25, Appl
18 1643.5 94.0 364 14 US-10-188-012-27 Sequence 27, Appl
19 1643.5 94.0 364 17 US-10-663-497-27 Sequence 27, Appl
20 1587 90.7 339 17 US-10-391-939A-2 Sequence 2, Appl
21 1587 90.7 339 17 US-10-391-939A-28 Sequence 28, Appl
22 1393 79.6 263 17 US-10-391-939A-6 Sequence 6, Appl
23 1393 79.6 263 17 US-10-391-939A-32 Sequence 32, Appl
24 1393 79.6 263 17 US-10-391-939A-39 Sequence 39, Appl
25 1393 79.6 263 17 US-10-391-939A-50 Sequence 50, Appl
26 592 33.8 108 17 US-10-805-177-50 Sequence 8, Appl
27 522.5 29.9 307 17 US-10-655-506-3 Sequence 3, Appl
28 519 29.7 305 14 US-10-188-012-1 Sequence 1, Appl
29 519 29.7 305 17 US-10-663-497-1 Sequence 1, Appl
30 491.5 28.1 282 14 US-10-188-012-3 Sequence 3, Appl
31 491.5 28.1 282 17 US-10-663-497-3 Sequence 3, Appl
32 432 24.7 81 17 US-10-718-321-6 Sequence 6, Appl
33 405 23.2 345 14 US-10-188-012-13 Sequence 13, Appl
34 405 23.2 345 17 US-10-663-497-13 Sequence 13, Appl
35 401 22.9 305 14 US-10-188-012-5 Sequence 5, Appl
36 401 22.9 305 14 US-10-188-012-7 Sequence 7, Appl
37 401 22.9 305 17 US-10-663-497-5 Sequence 5, Appl
38 401 22.9 305 17 US-10-663-497-7 Sequence 7, Appl
39 401 22.9 345 14 US-10-188-012-15 Sequence 15, Appl
40 401 22.9 345 17 US-10-663-497-15 Sequence 15, Appl
41 367.5 21.0 378 14 US-10-188-012-33 Sequence 33, Appl
42 367.5 21.0 378 17 US-10-663-497-33 Sequence 33, Appl
43 367.5 21.0 379 10 US-09-813-153-138 Sequence 138, App
44 366.5 21.0 378 14 US-10-188-012-35 Sequence 35, Appl
45 366.5 21.0 378 17 US-10-663-497-35 Sequence 35, Appl

ALIGNMENTS

RESULT 1

US-10-655-506-7
; Sequence 7, Application US/10655506
; Publication No. US20050089868A1
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Heesion
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/10/655,506
; FILING DATE: 04-Sep-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970
; FILING DATE: 23-Nov-1998
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996

ATTORNEY/AGENT INFORMATION:
NAME: Levine, Leslie M.
REGISTRATION NUMBER: 35,245
REFERENCE/DOCKET NUMBER: A010 PCT CIP-
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 679-2810
TELEFAX: (617) 679-2838
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 334 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-655-506-7

Query Match 100.0%; Score 1749; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
DB 1 MHPQVILSLHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
QY 61 IWWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IWWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPITVTTVTTRTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180
DB 121 SLEIVPPKVTTPITVTTVTTRTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180
QY 181 TTTSVPTTSTPTTTSVPVTTTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 240
DB 181 TTTSVPTTSTPTTTSVPVTTTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 240
QY 241 IRREPTSSPLSYTTDGNDDVTESDGLWNNNOTQLFLEHLLTANTTKGIYAGVCISVL 300
DB 241 IRREPTSSPLSYTTDGNDDVTESDGLWNNNOTQLFLEHLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIAKKYFFKKEVQQLRPHKSCIHQRE 334
DB 301 VLLALLGVIIAKKYFFKKEVQQLRPHKSCIHQRE 334

RESULT 2

US-10-718-321-7
Sequence 7, Application US/10718321
Publication No. US20050112117A1
GENERAL INFORMATION:
APPLICANT: Biogen Idec Ma Inc.
APPLICANT: Bailly, Veronique
APPLICANT: Bonventre, Joseph
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Molecules and Methods for Inhibiting
Shedding of KIM-1
FILE REFERENCE: A124 US
CURRENT APPLICATION NUMBER: US/10/718,321
CURRENT FILING DATE: 2003-11-20
PRIOR APPLICATION NUMBER: 60/295449
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: 60/295907
PRIOR FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: PCT/US02/17402
PRIOR FILING DATE: 2002-05-31
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 334
TYPE: PRT
ORGANISM: Homo Sapiens
US-10-718-321-7

Query Match 100.0%; Score 1749; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.5e-113;
Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
DB 1 MHPQVILSLHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
QY 61 IWWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IWWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPITVTTVTTRTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180
DB 121 SLEIVPPKVTTPITVTTVTTRTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180
QY 181 TTTSVPTTSTPTTTSVPVTTTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 240
DB 181 TTTSVPTTSTPTTTSVPVTTTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTTPT 240
QY 241 IRREPTSSPLSYTTDGNDDVTESDGLWNNNOTQLFLEHLLTANTTKGIYAGVCISVL 300
DB 241 IRREPTSSPLSYTTDGNDDVTESDGLWNNNOTQLFLEHLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIAKKYFFKKEVQQLRPHKSCIHQRE 334
DB 301 VLLALLGVIIAKKYFFKKEVQQLRPHKSCIHQRE 334

RESULT 3

US-10-188-012-17
Sequence 17, Application US/10188012
Publication No. US20030124114A1
GENERAL INFORMATION:
APPLICANT: McIntire, Jennifer Jones
APPLICANT: Umetsu, Dale T.
APPLICANT: Dekruyff, Rosemarie
APPLICANT: Kuchroo, Vijay
APPLICANT: Freeman, Gordon J.
TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
FILE REFERENCE: Use Thereof
CURRENT APPLICATION NUMBER: US/10/188,012
CURRENT FILING DATE: 2002-07-01
PRIOR APPLICATION NUMBER: 60/302,344
PRIOR FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 36
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 359
TYPE: PRT
ORGANISM: H. sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(360)
OTHER INFORMATION: TIM-1 allele 1
US-10-188-012-17

Query Match 96.3%; Score 1684; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
DB 1 MHPQVILSLHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
QY 61 IWWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IWWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPITVTTVTTRTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180
DB 121 SLEIVPPKVTTPITVTTVTTRTSTTTPVTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

```
QY 181 TTTSVPTTTSIPPTTSVPVTTTSTFVPPMPPLPRQNHEPVATSPSPQPAETHPTTLOGA 240
|
|
|
Db 181 TTTSVPTTTSIPPTTSVPVTTTSTFVPPMPPLPRQNHEPVATSPSPQPAETHPTTLOGA 240
|
|
|
QY 241 IRREPTSSPLSYTTDGNDDTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
|
|
|
Db 241 IRREPTSSPLSYTTDGNDDTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
|
|
|
QY 301 VLLALLGVIIIAKKYFFKKEVQQL 323
|
|
|
Db 301 VLLALLGVIIIAKKYFFKKEVQQL 323
|
|
|
```

```
RESULT 4
US-10-295-027-302
; Sequence 302, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 302
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-302
```

```
Query Match 96.3%; Score 1684; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
|
|
|
Db 1 MHPQVILSLIHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
|
|
|
QY 61 IWWTNGTHVTKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
|
|
|
```

```
Db 61 IWWTNGTHVTKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
|
|
|
QY 121 SLEIVPPKVTTPPIVTTVPTTTRTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTVS 180
|
|
|
Db 121 SLEIVPPKVTTPPIVTTVPTTTRTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTVS 180
|
|
|
QY 181 TTTSVPTTTSIPPTTSVPVTTTSTFVPPMPPLPRQNHEPVATSPSPQPAETHPTTLOGA 240
|
|
|
Db 181 TTTSVPTTTSIPPTTSVPVTTTSTFVPPMPPLPRQNHEPVATSPSPQPAETHPTTLOGA 240
|
|
|
QY 241 IRREPTSSPLSYTTDGNDDTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
|
|
|
Db 241 IRREPTSSPLSYTTDGNDDTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
|
|
|
QY 301 VLLALLGVIIIAKKYFFKKEVQQL 323
|
|
|
Db 301 VLLALLGVIIIAKKYFFKKEVQQL 323
|
|
|

RESULT 5
US-10-188-832-64
; Sequence 64, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; FILE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-64

Query Match 96.3%; Score 1684; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
|
|
|
Db 1 MHPQVILSLIHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCWNRGSCSLFTCQNG 60
|
|
|
QY 61 IWWTNGTHVTKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
|
|
|
Db 61 IWWTNGTHVTKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
|
|
|
QY 121 SLEIVPPKVTTPPIVTTVPTTTRTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTVS 180
|
|
|
Db 121 SLEIVPPKVTTPPIVTTVPTTTRTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTVS 180
|
|
|
QY 181 TTTSVPTTTSIPPTTSVPVTTTSTFVPPMPPLPRQNHEPVATSPSPQPAETHPTTLOGA 240
|
|
|
Db 181 TTTSVPTTTSIPPTTSVPVTTTSTFVPPMPPLPRQNHEPVATSPSPQPAETHPTTLOGA 240
|
|
|
QY 241 IRREPTSSPLSYTTDGNDDTVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
|
|
|
```

Db 241 IRREPTSSPLSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300
Qy 301 VLLALLGVIIIAKKYFFKKEVQOL 323
Db 301 VLLALLGVIIIAKKYFFKKEVQOL 323

RESULT 6

US-10-391-939A-4
; Sequence 4, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Giot, Loic;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Mesri, Mehdi;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glennda;
; APPLICANT: Starling, Gary;
; APPLICANT: Tse, Kam-Fai
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-048
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-4

Query Match 96.3%; Score 1684; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSCMNRGSCSLFTCQNG 60
Db 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSCMNRGSCSLFTCQNG 60
Qy 61 IVWTNGTHVTKDTRYKLLGDLSDRLDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Db 61 IVWTNGTHVTKDTRYKLLGDLSDRLDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Qy 121 SLEIVPPKVTTPITVTTVTVTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTT 180
Db 121 SLEIVPPKVTTPITVTTVTVTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTT 180
Qy 181 TTTSVPTTSTPTTTSVPVTTTSTVFVPPMPLPRQNHPEPVATSPSSQPAETHPTTLOGA 240
Db 181 TTTSVPTTSTPTTTSVPVTTTSTVFVPPMPLPRQNHPEPVATSPSSQPAETHPTTLOGA 240
Qy 241 IRREPTSSPLSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300
Db 241 IRREPTSSPLSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300
Qy 301 VLLALLGVIIIAKKYFFKKEVQOL 323
Db 301 VLLALLGVIIIAKKYFFKKEVQOL 323

RESULT 7

US-10-663-497-17
; Sequence 17, Application US/10663497
; Publication No. US2005009593A1
; GENERAL INFORMATION:

; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McInire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; FILE REFERENCE: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-663-497-17

Query Match 96.3%; Score 1684; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSCMNRGSCSLFTCQNG 60
Db 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSCMNRGSCSLFTCQNG 60
Qy 61 IVWTNGTHVTKDTRYKLLGDLSDRLDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Db 61 IVWTNGTHVTKDTRYKLLGDLSDRLDVSLLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Qy 121 SLEIVPPKVTTPITVTTVTVTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTT 180
Db 121 SLEIVPPKVTTPITVTTVTVTSTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTT 180
Qy 181 TTTSVPTTSTPTTTSVPVTTTSTVFVPPMPLPRQNHPEPVATSPSSQPAETHPTTLOGA 240
Db 181 TTTSVPTTSTPTTTSVPVTTTSTVFVPPMPLPRQNHPEPVATSPSSQPAETHPTTLOGA 240
Qy 241 IRREPTSSPLSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300
Db 241 IRREPTSSPLSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKGIYAGVCISVL 300
Qy 301 VLLALLGVIIIAKKYFFKKEVQOL 323
Db 301 VLLALLGVIIIAKKYFFKKEVQOL 323

RESULT 8

US-10-718-321-8
; Sequence 8, Application US/10718321
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE REFERENCE: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402


```
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-8

Query Match          96.3%; Score 1684; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
QY 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPITVTTVPTVTRTSTVPTTTTPTTTPPTTMSIPTTTTPTTMTVS 180
DB 121 SLEIVPPKVTTPITVTTVPTVTRTSTVPTTTTPTTTPPTTMSIPTTTTPTTMTVS 180
QY 181 TTTSVPTTTSIPTTTSVPVTTTSTVFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
DB 181 TTTSVPTTTSIPTTTSVPVTTTSTVFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
QY 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
DB 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIIAKKYFFKKEVQQL 323
DB 301 VLLALLGVIIIAKKYFFKKEVQQL 323

RESULT 9
US-10-847-918-25
; Sequence 25, Application US/10847918
; Publication No. US20050119210A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Be, Xiaobing
; APPLICANT: Liu, Wei
; APPLICANT: Slonim, Donna
; APPLICANT: Howes, Steve
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating Cancers
; FILE REFERENCE: 031896-026000 (AM101264)
; CURRENT APPLICATION NUMBER: US/10/847,918
; CURRENT FILING DATE: 2004-05-19
; PRIOR FILING DATE: 2004-05-19
; PRIOR FILING DATE: 2003-05-20
; NUMBER OF SEQ ID NOS: 14937
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-847-918-25

Query Match          96.3%; Score 1684; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 5.2e-109;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
QY 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120

Query Match          96.0%; Score 1679; DB 14; Length 359;
Best Local Similarity 99.7%; Pred. No. 1.2e-108;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
QY 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPITVTTVPTVTRTSTVPTTTTPTTTPPTTMSIPTTTTPTTMTVS 180
DB 121 SLEIVPPKVTTPITVTTVPTVTRTSTVPTTTTPTTTPPTTMSIPTTTTPTTMTVS 180
QY 181 TTTSVPTTTSIPTTTSVPVTTTSTVFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
DB 181 TTTSVPTTTSIPTTTSVPVTTTSTVFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
QY 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
DB 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIIAKKYFFKKEVQQL 323
DB 301 VLLALLGVIIIAKKYFFKKEVQQL 323

RESULT 10
US-10-188-012-19
; Sequence 19, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-188-012-19

Query Match          96.0%; Score 1679; DB 14; Length 359;
Best Local Similarity 99.7%; Pred. No. 1.2e-108;
Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCNWNGSCSLFTCQNG 60
QY 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IVWTNGTHVTRKDYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPITVTTVPTVTRTSTVPTTTTPTTTPPTTMSIPTTTTPTTMTVS 180
DB 121 SLEIVPPKVTTPITVTTVPTVTRTSTVPTTTTPTTTPPTTMSIPTTTTPTTMTVS 180
QY 181 TTTSVPTTTSIPTTTSVPVTTTSTVFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
DB 181 TTTSVPTTTSIPTTTSVPVTTTSTVFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
QY 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
DB 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIIAKKYFFKKEVQQL 323
DB 301 VLLALLGVIIIAKKYFFKKEVQQL 323
```

Db 301 VLLALLGVIIIAKKYFFKKEVQOL 323

RESULT 11

US-10-663-497-19

; Sequence 19, Application US/10663497

; Publication No. US20050095593A1

; GENERAL INFORMATION:

; APPLICANT: Dale Umetsu

; APPLICANT: Rosemarie DeKruyff

; APPLICANT: Jennifer McIntire

; APPLICANT: Gordon Freeman

; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH

; TITLE OF INVENTION: IMMUNE DISEASE

; FILE REFERENCE: STAN-235CIP

; CURRENT APPLICATION NUMBER: US/10/663,497

; CURRENT FILING DATE: 2003-09-15

; PRIOR APPLICATION NUMBER: 60/302,344

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 10/188,012

; PRIOR FILING DATE: 2002-07-01

; NUMBER OF SEQ ID NOS: 40

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 19

; LENGTH: 359

; TYPE: PRT

; ORGANISM: H. sapiens

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(359)

; OTHER INFORMATION: TIM-1, allele 2

US-10-663-497-19

Query Match 96.0%; Score 1679; DB 17; Length 359;

Best Local Similarity 99.7%; Pred. No. 1.2e-108;

Matches 322; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60

Db 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60

QY 61 IVWTNGTHVTRKDTRYKLLGDLRRDVSILTIENTAVSDSGVYCCRVHRGWFNMDKITV 120

Db 61 IVWTNGTHVTRKDTRYKLLGDLRRDVSILTIENTAVSDSGVYCCRVHRGWFNMDKITV 120

QY 121 SLEIVPPKVTTPITVTTVTRTSTVTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

Db 121 SLEIVPPKVTTPITVTTVTRTSTVTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

QY 181 TTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTT 240

Db 181 TTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTT 240

QY 241 IRREPTSSPLSYTTDGNPTVTSDDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

Db 241 IRREPTSSPLSYTTDGNPTVTSDDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

QY 301 VLLALLGVIIIAKKYFFKKEVQOL 323

Db 301 VLLALLGVIIIAKKYFFKKEVQOL 323

RESULT 12

US-10-188-012-23

; Sequence 23, Application US/10188012

; Publication No. US20030124114A1

; GENERAL INFORMATION:

; APPLICANT: McIntire, Jennifer Jones

; APPLICANT: Umetsu, Dale T.

; APPLICANT: DeKruyff, Rosemarie

; APPLICANT: Kuchroo, Vijay

; APPLICANT: Freeman, Gordon J.

; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of

; TITLE OF INVENTION: Use Thereof

; FILE REFERENCE: STAN-235

; CURRENT APPLICATION NUMBER: US/10/188,012

; CURRENT FILING DATE: 2002-07-01

; PRIOR APPLICATION NUMBER: 60/302,344

; PRIOR FILING DATE: 2001-06-29

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 23

; LENGTH: 359

; TYPE: PRT

; ORGANISM: H. sapiens

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(359)

; OTHER INFORMATION: TIM-1, allele 4

US-10-188-012-23

Query Match 95.5%; Score 1671; DB 14; Length 359;

Best Local Similarity 99.1%; Pred. No. 4.2e-108;

Matches 320; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60

Db 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60

QY 61 IVWTNGTHVTRKDTRYKLLGDLRRDVSILTIENTAVSDSGVYCCRVHRGWFNMDKITV 120

Db 61 IVWTNGTHVTRKDTRYKLLGDLRRDVSILTIENTAVSDSGVYCCRVHRGWFNMDKITV 120

QY 121 SLEIVPPKVTTPITVTTVTRTSTVTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

Db 121 SLEIVPPKVTTPITVTTVTRTSTVTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

QY 181 TTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTT 240

Db 181 TTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTT 240

QY 241 IRREPTSSPLSYTTDGNPTVTSDDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

Db 241 IRREPTSSPLSYTTDGNPTVTSDDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300

QY 301 VLLALLGVIIIAKKYFFKKEVQOL 323

Db 301 VLLALLGVIIIAKKYFFKKEVQOL 323

RESULT 13

US-10-663-497-23

; Sequence 23, Application US/10663497

; Publication No. US20050095593A1

; GENERAL INFORMATION:

; APPLICANT: Dale Umetsu

; APPLICANT: Rosemarie DeKruyff

; APPLICANT: Jennifer McIntire

; APPLICANT: Gordon Freeman

; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH

; TITLE OF INVENTION: IMMUNE DISEASE

; FILE REFERENCE: STAN-235CIP

; CURRENT APPLICATION NUMBER: US/10/663,497

; CURRENT FILING DATE: 2003-09-15

; PRIOR APPLICATION NUMBER: 60/302,344

; PRIOR FILING DATE: 2001-06-29

; PRIOR APPLICATION NUMBER: 10/188,012

; PRIOR FILING DATE: 2002-07-01

; NUMBER OF SEQ ID NOS: 40

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 23

; LENGTH: 359

; TYPE: PRT

; ORGANISM: H. sapiens

; FEATURE:

; NAME/KEY: VARIANT

```
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 4
US-10-663-497-23

Query Match      95.5%; Score 1671; DB 17; Length 359;
Best Local Similarity 99.1%; Pred. No. 4.2e-108;
Matches 320; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCCHYSGAVTSMCNWRGSCSLFTCCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCCHYSGAVTSMCNWRGSCSLFTCCQNG 60
QY 61 IWWTNGTHVTKDTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IWWTNGTHVTKDTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPPIVTTVPVTTVTRTSTVPTTTTTPVTTTTPVTTTTPVTTTTPV 180
DB 121 SLEIVPPKVTTPPIVTTVPVTTVTRTSTVPTTTTTPVTTTTPVTTTTPVTTTTPV 180
QY 181 TTTSVPTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPTTLOQA 240
DB 181 TTTSVPTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPTTLOQA 240
QY 241 IRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
DB 241 IRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIAKKYFFKKEVQQL 323
DB 301 VLLALLGVIIAKKYFFKKEVQQL 323

RESULT 14
US-10-188-012-21
; Sequence 21, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188, 012
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 365
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(365)
; OTHER INFORMATION: TIM-1, allele 3
US-10-188-012-21

Query Match      95.3%; Score 1666; DB 14; Length 365;
Best Local Similarity 97.9%; Pred. No. 9.4e-108;
Matches 322; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCCHYSGAVTSMCNWRGSCSLFTCCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCCHYSGAVTSMCNWRGSCSLFTCCQNG 60
QY 61 IWWTNGTHVTKDTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IWWTNGTHVTKDTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPPIVTTVPVTTVTRTSTVPTTTTTPVTTTTPVTTTTPVTTTTPV 180
DB 121 SLEIVPPKVTTPPIVTTVPVTTVTRTSTVPTTTTTPVTTTTPVTTTTPVTTTTPV 180
QY 175 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 234
DB 175 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 234
QY 181 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 240
DB 181 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 240
QY 235 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 294
DB 235 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 294
QY 241 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 300
DB 241 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 300
QY 295 VCISVLVLLALLGVIIAKKYFFKKEVQQL 323
DB 295 VCISVLVLLALLGVIIAKKYFFKKEVQQL 323

US-10-663-497-21
; Sequence 21, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie Dekruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 365
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(365)
; OTHER INFORMATION: TIM-1, allele 3
US-10-663-497-21

Query Match      95.3%; Score 1666; DB 17; Length 365;
Best Local Similarity 97.9%; Pred. No. 9.4e-108;
Matches 322; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCCHYSGAVTSMCNWRGSCSLFTCCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCCHYSGAVTSMCNWRGSCSLFTCCQNG 60
QY 61 IWWTNGTHVTKDTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
DB 61 IWWTNGTHVTKDTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPPIVTTVPVTTVTRTSTVPTTTTTPVTTTTPVTTTTPVTTTTPV 174
DB 121 SLEIVPPKVTTPPIVTTVPVTTVTRTSTVPTTTTTPVTTTTPVTTTTPVTTTTPV 180
QY 175 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 234
DB 175 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 234
QY 181 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 240
DB 181 TTMTVSTTTSVPVTTTTSIPTTTSVPVTTTVPVPPMPLPRQNHPEVATSPSSPQPAETHPT 240
QY 235 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 294
DB 235 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 294
QY 241 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 300
DB 241 TTLOGAIRREPTSSPLSYTTDNDVTVESSDGLMNNQTLFLEHSLLTANTTKGIYAG 300
QY 295 VCISVLVLLALLGVIIAKKYFFKKEVQQL 323
DB 295 VCISVLVLLALLGVIIAKKYFFKKEVQQL 323
```

Db 301 VCISVLVLLALLGVIIAKKYFFKKEVQOL 329

Search completed: June 29, 2005, 09:33:01
Job time : 210.324 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 45.9332 Seconds
(without alignments)
542.805 Million cell updates/sec

Title: US-10-718-321-7

Perfect score: 1749

Sequence: 1 MHPQVILSLHLADSVAG.....FFKKEVQQLRPHKSHQRE 334

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB.pcp:*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB.pcp:*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB.pcp:*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB.pcp:*
- 5: /cgn2_6/ptodata/1/1aa/PTUS COMB.pcp:*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1749	100.0	334	4	US-09-197-970B-7
2	1309	74.8	451	1	US-08-287-001A-2
3	1309	74.8	451	5	PCT-US95-09941-2
4	522.5	29.9	307	4	US-09-197-970B-3
5	292	16.7	281	1	US-08-487-748A-9
6	292	16.7	281	1	US-08-487-748A-10
7	292	16.7	281	3	US-08-480-070C-10
8	292	16.7	281	3	US-08-829-525-10
9	292	16.7	281	3	US-08-609-583A-10
10	292	16.7	281	3	US-08-937-399-10
11	292	16.7	281	4	US-09-310-367-10
12	292	16.7	281	4	US-09-032-337-10
13	292	16.7	281	4	US-09-464-231-10
14	267	15.3	301	3	US-08-829-525-24
15	267	15.3	301	3	US-08-609-583A-24
16	267	15.3	301	3	US-08-937-399-24
17	267	15.3	301	4	US-09-310-367-24
18	267	15.3	301	4	US-09-032-337-24
19	267	15.3	301	4	US-09-464-231-24
20	235	13.4	786	3	US-09-103-429A-3
21	235	13.4	805	3	US-09-103-429A-4
22	232.5	13.3	288	4	US-09-216-393B-341
23	232.5	13.3	288	4	US-09-216-393B-344
24	229	13.1	788	4	US-09-294-663-3
25	229	13.1	807	4	US-09-294-663-4
26	227.5	13.0	1721	3	US-08-700-651-5
27	227.5	13.0	1721	3	US-08-928-361B-6

28	227.5	13.0	1721	4	US-09-588-995A-6	Sequence 6, Appli
29	227.5	13.0	1837	3	US-08-928-361B-5	Sequence 5, Appli
30	227.5	13.0	1837	4	US-09-588-995A-5	Sequence 5, Appli
31	226.5	13.0	5179	4	US-09-538-092-1258	Sequence 1258, Ap
32	225	12.9	216	3	US-08-928-361B-27	Sequence 27, Appl
33	222	12.7	216	3	US-08-928-361B-8	Sequence 8, Appli
34	222	12.7	216	4	US-09-588-995A-8	Sequence 8, Appli
35	221.5	12.7	175	3	US-08-700-651-12	Sequence 12, Appl
36	221.5	12.7	175	3	US-08-928-361B-17	Sequence 17, Appl
37	221.5	12.7	175	4	US-09-588-995A-17	Sequence 17, Appl
38	219.5	12.6	150	3	US-08-928-361B-18	Sequence 18, Appl
39	219.5	12.6	150	4	US-09-588-995A-18	Sequence 18, Appl
40	219.5	12.6	249	3	US-08-700-651-15	Sequence 15, Appl
41	219.5	12.6	249	3	US-08-928-361B-20	Sequence 20, Appl
42	219.5	12.6	249	4	US-09-588-995A-20	Sequence 20, Appl
43	208.5	11.9	138	3	US-08-700-651-10	Sequence 10, Appl
44	208.5	11.9	138	3	US-08-928-361B-15	Sequence 15, Appl
45	208.5	11.9	138	4	US-09-588-995A-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1

US-09-197-970B-7

; Sequence 7, Application US/09197970B

; Patent No. 6664385

; GENERAL INFORMATION:

APPLICANT: Michele Sanicola-Nadel

Joseph V. Bonventre

Catherine A. Hession

Takaharu Ichimura

Henry Wei

Richard L. Cate

TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Biogen, Inc.

STREET: 14 Cambridge Center

CITY: Cambridge

STATE: MA

COUNTRY: USA

ZIP: 02142

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/197,970B

FILING DATE: 23-No. 6664385-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/018,228

FILING DATE: 24-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Levine, Leslie M.

REGISTRATION NUMBER: 35,245

REFERENCE/DOCKET NUMBER: A010 PCT CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 679-2810

TELEFAX: (617) 679-2838

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 334 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-197-970B-7

Query Match

Best Local Similarity 100.0%; Score 1749; DB 4; Length 334;

100.0%; Pred. No. 7.9e-136;

Best Local Similarity 63.3%; Pred. No. 1.6e-99;
Matches 255; Conservative 36; Mismatches 32; Indels 80; Gaps 3;
QY 1 MHPOVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCMNRGSCSLFTCONG 60
DB 1 MHLQVILSLIHLADSVADSVNVDGAGLSITLPCRYNGAITSMCNNGTCSVSCPDG 60
QY 61 IWTNGTHVTRKRYKLLGLDLSRRDVSLLTIENTAVSDSGVYCCRVHRGWFNDKMTIV 120
DB 61 IWTNGTHVTRKRYKLLGLNLSRRDVSLLTIENTAVSDSGIYCCRVKXHSWFNDKMTI 120
QY 121 SLEIVPPKV-----TTTTPIVTVTP 139
DB 121 SLEIGPPRVTPIVTRFTSTVPTTTTLPPTTTTLPPTTTTLPMTTLPMTTLP 180
QY 140 TVTTRTSTVPTTTVP-----TTTVPPTMSIPTTTTVPPTMTVSTT 182
DB 181 TTTTVPPTTTTLPPTTLPMTTLPPTTLPPTTLPPTTLPMTTLPMTTLPPTTLPPT 240
QY 183 T-----SVPTTTSIPTTSVPVTTTSTVTFVPPMPLPRQNHBPV 220
DB 241 TLPPTTLPPTTLPPTTLPMTTLPPTTLPPTTLPPTTLPMTTLPMTVSTFVPPTP 300
QY 221 ATSPSSPOPAETHPTTLOGAIRREPTSSPLSYTTDGNVTSTESSDGLWNNQTLFLEH 280
DB 301 ATSPSSPOPAETHPTTLLGATRTQTSSPLSYTTDGSVTSTESSDGLWNNQTLSP 360
QY 281 SLLTANTTKGIYAGVCSIVLLALLGLVIAKKYFPKKEVQQL 323
DB 361 SPQWNTTEGIYAGVCSIVLLALLGLVIAKKYFPKKEIQQL 403

RESULT 4
US-09-197-970B-3
; Sequence 3, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-NOV-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid

LENGTH: 307 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-197-970B-3
Query Match 29.9%; Score 522.5; DB 4; Length 307;
Best Local Similarity 39.4%; Pred. No. 4.2e-35;
Matches 128; Conservative 39; Mismatches 93; Indels 65; Gaps 12;
QY 4 QVVILSLIHLADSVAGSVKVGEGAPSVTLPCYSGAVTSMCMNRGSCSLFTCONGI 61
DB 5 QVIFGLLLPLPGSVDSYEVWVGVPVTPCTYSTRGIIITTCWGRGQCQPYSSCONIL 64
QY 62 VMTNGTHVTRKRYKLLGLDLSRRDVSLLTIENTAVSDSGVYCCRVHRGWFNDKMTIVS 121
DB 65 IWTNGVQVYTRSSGRYNIKRISEGDVSLTIENSVDSDGLYCCRVHPGFNDQKMTFS 124
QY 122 LEIVPPKVTTPIVTVTVTPTTSTVPT-----TTTTVPTTVPPTMSIPTTTTVPPTMTV 179
DB 125 LE-VKPEIPTSP-----PTRPTTTRPTTPTTIS----- 153
QY 180 STTTSVPPTTSTPTTSTVPTTSTVTFVPPMPLPRQNHBPVATSPSSPOPAETHPTTLOG 239
DB 154 TRSTHPTSTRVSTST-----PTPEQTQTHKPEITT-----FYAHETT----- 191
QY 240 AIRREPTSSPLSYTTDGNVTSTESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCSIV 299
DB 192 ---AEVTETPSYT-PADWNGIVT-SSEAWNNHTVRIPLRKP--QRNPTKGFVVGMSVAA 244
QY 300 LVLLALLGLVIAKKY-FPKKEVQQL 323
DB 245 LLLLLLASTVTVVTVYIIIRKMGSL 269
RESULT 5
US-08-487-748A-9
; Sequence 9, Application US/08487748A
; Patent No. 5721351
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,748A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid

TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-487-748A-9

Query Match 16.7%; Score 292; DB 1; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VILSLIHLADSVAGSVKVGGEAGPSVTLPCY-----SGAVTSMCMWRGSCSLFTCQNGI 61
DB 10 VLLLLQLLARSLELDGYK--EVGNAYLPCSYTLPTSGLVPMCGWKGFCFWSQCTNEL 67
QY 62 VWTNGTHVYKOTRYKLLGDLRRDVSLLTENTAVSDSGVYCCRVHRGWFNDMKITVS 121
DB 68 LRTDERNVYQKSRYYQLKGLNKGDSLIINKVTLDDHGYCCRIQFPGLMNDKKLELK 127
QY 122 LEIVPPKVTTPPIVTTVPTVTTVTSITVPTTTTPTTTTPTTTTPTTTTPTTTTPTT 181
DB 128 LDIAKAVT-----PRT----- 136
QY 182 TTSVPTTTSIPTTISVPVTTTTSVTFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGAI 241
DB 137 -----PAQTAHGDSITTA-----PRT----- 153
QY 242 RREPTSSPLSYTTDGNVTTESSDGLWNNNOTOLF-----LEHSLLTANTTKGIYAGVC 296
DB 154 -----TTERNGSETQTLVTLHNNNGTKISTWADIKDSGETIRT--AIHIGVG 199
QY 297 ISV-LVLLALGLVIAKKYFPKKE 319
DB 200 VSAGLTALIIIGVLILKWYSCKKK 223

RESULT 6
US-08-487-748A-10
; Sequence 10, Application US/08487748A
; Patent No. 5721351
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,748A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-487-748A-10

Query Match 16.7%; Score 292; DB 1; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VILSLIHLADSVAGSVKVGGEAGPSVTLPCY-----SGAVTSMCMWRGSCSLFTCQNGI 61
DB 10 VLLLLQLLARSLELDGYK--EVGNAYLPCSYTLPTSGLVPMCGWKGFCFWSQCTNEL 67
QY 62 VWTNGTHVYKOTRYKLLGDLRRDVSLLTENTAVSDSGVYCCRVHRGWFNDMKITVS 121
DB 68 LRTDERNVYQKSRYYQLKGLNKGDSLIINKVTLDDHGYCCRIQFPGLMNDKKLELK 127
QY 122 LEIVPPKVTTPPIVTTVPTVTTVTSITVPTTTTPTTTTPTTTTPTTTTPTTTTPTT 181
DB 128 LDIAKAVT-----PRT----- 136
QY 182 TTSVPTTTSIPTTISVPVTTTTSVTFVPPMPLPRQNHPEVATSPSPQPAETHPTTLOGAI 241
DB 137 -----PAQTAHGDSITTA-----PRT----- 153
QY 242 RREPTSSPLSYTTDGNVTTESSDGLWNNNOTOLF-----LEHSLLTANTTKGIYAGVC 296
DB 154 -----TTERNGSETQTLVTLHNNNGTKISTWADIKDSGETIRT--AIHIGVG 199
QY 297 ISV-LVLLALGLVIAKKYFPKKE 319
DB 200 VSAGLTALIIIGVLILKWYSCKKK 223

RESULT 7
US-08-480-070C-10
; Sequence 10, Application US/08480070C
; Patent No. 6066498
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,070C
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-480-070C-10

Query Match 16.7%; Score 292; DB 3; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;

Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VILSLILHLADSVAGSKVGEAGPSVTLPCYH-----SGAVTSMCNRGSCSLFTQNGI 61
DB 10 VLLLLQLLLARSLEDEGKVV--EVGNAYLPCS YTLPTSGTLVPMCKGKFCPWSQCTNEL 67
QY 62 VWTNGTHVYRKDTRKYLGLDLSRRDVS LTIENAVSDSGVYCCRVHRGWFNMDKITVS 121
DB 68 LRTDERNVYQKSSRYQLKGLDNLKGDVSLIKNVTLDDHGTCCRIQFPGMLMNDKKLELX 127
QY 122 LEIVPPKVTTPITVTTVPTVTTVTSTTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
DB 128 LDIAKAVT-----PRTLL----- 136
QY 182 TTSVPTTTSIPTTSVPVTTTSTFVPPMPLPRQNHVPATSPSPQPAETHPTTLQGA 241
DB 137 -----PAQTAHGDSTTAS-----PRTLL----- 153
QY 242 RREPTSSPLYSTTGDNDVTSSDGLWNNQTLF-----LEHSLLTANTTKGIYAGVC 296
DB 154 -----TTERNGSETQTLVTLHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199
QY 297 ISV-LVLLALLGVIIAKYFFKKE 319
DB 200 VSAGLTALIIGVLIILKWKYSCKK 223

RESULT 8

US-08-829-525-10
; Sequence 10, Application US/08829525
; Patent No. 6084083
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,525
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-081
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein

US-08-829-525-10

Query Match 16.7%; Score 292; DB 3; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VILSLILHLADSVAGSKVGEAGPSVTLPCYH-----SGAVTSMCNRGSCSLFTQNGI 61
DB 10 VLLLLQLLLARSLEDEGKVV--EVGNAYLPCS YTLPTSGTLVPMCKGKFCPWSQCTNEL 67
QY 62 VWTNGTHVYRKDTRKYLGLDLSRRDVS LTIENAVSDSGVYCCRVHRGWFNMDKITVS 121
DB 68 LRTDERNVYQKSSRYQLKGLDNLKGDVSLIKNVTLDDHGTCCRIQFPGMLMNDKKLELX 127
QY 122 LEIVPPKVTTPITVTTVPTVTTVTSTTTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
DB 128 LDIAKAVT-----PRTLL----- 136
QY 182 TTSVPTTTSIPTTSVPVTTTSTFVPPMPLPRQNHVPATSPSPQPAETHPTTLQGA 241
DB 137 -----PAQTAHGDSTTAS-----PRTLL----- 153
QY 242 RREPTSSPLYSTTGDNDVTSSDGLWNNQTLF-----LEHSLLTANTTKGIYAGVC 296
DB 154 -----TTERNGSETQTLVTLHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199
QY 297 ISV-LVLLALLGVIIAKYFFKKE 319
DB 200 VSAGLTALIIGVLIILKWKYSCKK 223

RESULT 9

US-08-609-583A-10
; Sequence 10, Application US/08609583A
; Patent No. 6204371
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,583A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid

```
;
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-609-583A-10

Query Match      16.7%; Score 292; DB 3; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VILSLILHLADSVAGSVKVGEGAGPSVTLPCY-----SGAVTSMCWNRGSCSLFTCONGI 61
Db 10 VLLQLQLLARSLEDEYKV--EVGNAYLPCSYTLPTSGTLVPMCWGKGFCEWFSQCTNEL 67
QY 62 VMTNGTHVYRKDTRYKLLGDLSDRRDVSLLTIENTAVSDGVCYCCVHRHGFNDMKITVS 121
Db 68 LRTDERNVYKSSRYQLKGLNKGDSLLIKNVLDHGYCCRIQPPGLMNDKKLELK 127
QY 122 LEIVPPKVTTPITVTVPTVTRTSTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
Db 128 LDIAKAAKVT-----PRTL----- 153
QY 182 TTSVPTTTSIPTTSVPVTVTVTFVPPMPLPRQNHPEVATSPSSPQPAETHPTTLOGAI 241
Db 137 -----PAQTAHGDSTTAS----- 153
QY 242 RREPTSPLYSYTDDGNDTVTESDGLMNNNOTOLF-----LEHSLLTANTTKGIYAGVC 296
Db 154 -----TTERNGSETQTLVTLHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199

QY 297 ISV-LVLLALLGVIIAKKYFPKKE 319
Db 200 VSAGLTALIIIGVILKWKYSCKK 223

RESULT 10
US-08-937-399-10
; Sequence 10, Application US/08937399
; Patent No. 6288218
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/937,399
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE

;
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-937-399-10

Query Match      16.7%; Score 292; DB 3; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VILSLILHLADSVAGSVKVGEGAGPSVTLPCY-----SGAVTSMCWNRGSCSLFTCONGI 61
Db 10 VLLQLQLLARSLEDEYKV--EVGNAYLPCSYTLPTSGTLVPMCWGKGFCEWFSQCTNEL 67
QY 62 VMTNGTHVYRKDTRYKLLGDLSDRRDVSLLTIENTAVSDGVCYCCVHRHGFNDMKITVS 121
Db 68 LRTDERNVYKSSRYQLKGLNKGDSLLIKNVLDHGYCCRIQPPGLMNDKKLELK 127
QY 122 LEIVPPKVTTPITVTVPTVTRTSTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
Db 128 LDIAKAAKVT-----PRTL----- 153
QY 182 TTSVPTTTSIPTTSVPVTVTVTFVPPMPLPRQNHPEVATSPSSPQPAETHPTTLOGAI 241
Db 137 -----PAQTAHGDSTTAS----- 153
QY 242 RREPTSPLYSYTDDGNDTVTESDGLMNNNOTOLF-----LEHSLLTANTTKGIYAGVC 296
Db 154 -----TTERNGSETQTLVTLHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199

QY 297 ISV-LVLLALLGVIIAKKYFPKKE 319
Db 200 VSAGLTALIIIGVILKWKYSCKK 223

RESULT 11
US-09-310-367-10
; Sequence 10, Application US/09310367
; Patent No. 6414117
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/310,367
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/829,525
; FILING DATE: 28-MAR-1997
; APPLICATION NUMBER: US 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
```

```
;
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-310-367-10

Query Match      16.7%; Score 292; DB 4; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VLSLILHLADSVAGSVKVGGEAGPSVTLPCHY-----SGAVTSMCWNRGSCSLFTCONGI 61
DB 10 VLLLLQLLARSLEDEGYKV--EVGNAYLPCSYTLPTSGTLVPMCMWKGKFCPWSQCTNEL 67
QY 62 VMTNGTHVTRKDYKLLGDLRRDVSILTIENTAVSDSGVYCCRVHRGWFNDMKITVS 121
DB 68 LRTDERNTYKSSRYQLKGLDNLKGDVSLIKNVLDHGTCCRIQPPGLMNDKKLELK 127
QY 122 LEIVPEKVTPIVTTVPTVTTVTRTSTTPTTTTPTTPTTPTTPTTPTTPTTPTTPTT 181
DB 128 LDIAKAKVT----- 136
QY 182 TTSVPTTTSIPPTTSVPVTTTSTVTPVPMPLPRQNHPEVATSPSPQPAETHPTTLOGAI 241
DB 137 -----PAQTAHGDSTTAS-----PRTL----- 153
QY 242 RREPTSSPLSYVTGNDTVTSSDGLWNNNOTLPF-----LEHSLLTANTTKGIYAGVC 296
DB 154 -----TTERNGSETQTLVTLHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199
QY 297 ISV-LVLLALLGVIIAKKYFFKKE 319
DB 200 VSAGLTALIIIGVLILKWYSCKKK 223

RESULT 12
US-09-337-10
; Sequence 10, Application US/09032337
; Patent No. 6455685
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032.337
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995

;
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-016
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-032-337-10

Query Match      16.7%; Score 292; DB 4; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

QY 6 VLSLILHLADSVAGSVKVGGEAGPSVTLPCHY-----SGAVTSMCWNRGSCSLFTCONGI 61
DB 10 VLLLLQLLARSLEDEGYKV--EVGNAYLPCSYTLPTSGTLVPMCMWKGKFCPWSQCTNEL 67
QY 62 VMTNGTHVTRKDYKLLGDLRRDVSILTIENTAVSDSGVYCCRVHRGWFNDMKITVS 121
DB 68 LRTDERNTYKSSRYQLKGLDNLKGDVSLIKNVLDHGTCCRIQPPGLMNDKKLELK 127
QY 122 LEIVPEKVTPIVTTVPTVTTVTRTSTTPTTTTPTTPTTPTTPTTPTTPTTPTTPTT 181
DB 128 LDIAKAKVT----- 136
QY 182 TTSVPTTTSIPPTTSVPVTTTSTVTPVPMPLPRQNHPEVATSPSPQPAETHPTTLOGAI 241
DB 137 -----PAQTAHGDSTTAS-----PRTL----- 153
QY 242 RREPTSSPLSYVTGNDTVTSSDGLWNNNOTLPF-----LEHSLLTANTTKGIYAGVC 296
DB 154 -----TTERNGSETQTLVTLHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199
QY 297 ISV-LVLLALLGVIIAKKYFFKKE 319
DB 200 VSAGLTALIIIGVLILKWYSCKKK 223

RESULT 13
US-09-464-231-10
; Sequence 10, Application US/09464231
; Patent No. 6562343
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/464,231
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 281 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-464-231-10

Query Match 16.7%; Score 292; DB 4; Length 281;
Best Local Similarity 27.5%; Pred. No. 3.2e-16;
Matches 89; Conservative 37; Mismatches 78; Indels 120; Gaps 8;

Qy 6 VILSLILHLADSVAGSKVGEAGPSVTLPCHY-----SGAVTSMCWNRGSCSLFTCQNGI 61
Db 10 VILLQLLLARSLEGGYKV--EVGNAYLPCSYTLFTSGTILVPMCWGKGFCFWSQCTNEL 67
Qy 62 VWTNGTHVYRKTRYKLLGDLSDRRDVSLLTIENTAVSDSGVYCCRVHRGNFNDMKITVS 121
Db 68 LRTDERNVYQKSRVQLKDLNGKDVSLIINKVTLDDHGYCCRIQPPGLMNDKLELK 127
Qy 122 LEIVPPKVTTPPIVTVPTVTRTSTVPTTTTPTVPTTPTVPTTPTVPTTPTVPTTPTVPT 181
Db 128 LDIKAAKVT-----PAQTAHGDSFTAS-----PRTL----- 153
Qy 182 TTSVPTTTSIPTTISVPVTTTSTVFPMPPLPRQNHEPVATSPSSQPAETHPTTLOAI 241
Db 137 -----PAQTAHGDSFTAS-----PRTL----- 153
Qy 242 RREPTSSPLYSYTTDGNVTVTSSDGLMNNNTQLF-----LEHSLLTANTTKGIYAGVC 296
Db 154 -----TTERNGSETQLTLVHNNNGTKISTWADEIKDSGETIRT--AIHIGVG 199
Qy 297 ISV-LVLLALLGVIIAKYFFKKE 319
Db 200 VSAGLTALIIIGVLILKWYSCKK 223

RESULT 14
US-08-829-525-24
; Sequence 24, Application US/08829525
; Patent No. 6084083
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/829,525
; FILING DATE: 28-MAR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/609,583
; FILING DATE: 01-MAR-1996
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 301 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-829-525-24

Query Match 15.3%; Score 267; DB 3; Length 301;
Best Local Similarity 26.9%; Pred. No. 4e-14;
Matches 88; Conservative 32; Mismatches 99; Indels 108; Gaps 11;

Qy 6 VILSLILHLADSVAGSKVGEAGPSVTLPCHY-----GAVTSMCWNRGSCSLFTCQNGI 61
Db 10 VILLLLLLLSTRS--SEVEYRAEVGNAYLPCFYTPAAPGNLVPVCWKGACPVFCGNV 67
Qy 62 VWTNGTHVYRKTRYKLLGDLSDRRDVSLLTIENTAVSDSGVYCCRVHRGNFNDMKITVS 121
Db 68 LRTDERDNYV-WTSRYWNLGDFRKGDVSLTIENVTLADSGIYCCRIQIPGIMNDEKFNK 126
Qy 122 LEIVPPKVTTPPIVTVPTVTRTSTVPTTTTPTVPTTPTVPTTPTVPTTPTVPTTPTVPT 181
Db 127 LVIPKAVKTPAP-----TLQORDFTAAPPRMLTT-- 154
Qy 182 TTSVPTTTSIPTTISVPVTTTSTVFPMPPLPRQNHEPVATSP--SSQPAETHPTTLOQ 239
Db 155 -----RHGHPAETQTLGSLPDINLTQISTLAN 181
Qy 240 AIRREPTSSPLYSYTTDGNVTVTSSDGLMNNNTQLFLEHSLLTANTTKGIY--AGVCI 297
Db 182 ELR-----DSRLANDLRDSGATI-----RIGIYIGAGICA 211
Qy 298 SVLVLLALLGVIIAKYFFKKE-VQOL 323
Db 212 G-LALALIFGALIFKWYSHSKEKIQL 237

RESULT 15
US-08-609-583A-24
; Sequence 24, Application US/08609583A
; Patent No. 6204371
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
```

Search completed: June 29, 2005, 09:02:59
Job time : 47.9332 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 237.54 Seconds
(without alignments)
543.815 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLHLADSVAG.....FFKKEVQQLRPHKSHQRE 334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1749	100.0	334	2	AAW38336 Human kid
2	1749	100.0	334	6	AAO26679 KIM-1 rel
3	1749	100.0	334	8	ADQ76696 Human kid
4	1684	96.3	359	6	ABR58582 Human can
5	1684	96.3	359	6	ABP70438 Amino aci
6	1684	96.3	359	6	ABR48174 Human bia
7	1684	96.3	359	6	AAO26680 KIM-1 rel
8	1684	96.3	359	7	ADN36594 Human NOV
9	1684	96.3	359	7	ADN38994 Cancer/an
10	1684	96.3	359	8	ADQ76690 Human kid
11	1679	96.0	359	6	ABP70439 Amino aci
12	1671	95.5	359	6	ABP70441 Amino aci
13	1666	95.3	365	6	ABP70440 Amino aci
14	1655.5	94.7	364	6	ABP70442 Amino aci
15	1643.5	94.0	364	6	ABP70443 Amino aci
16	1587	90.7	339	7	ADN36592 Plasmid p
17	1587	90.7	339	7	ADN36593 Human NOV
18	1529	87.4	298	8	ADQ76693 Human KIM
19	1529	87.4	518	8	ADQ76691 Human KIM
20	1393	79.6	263	7	ADN36629 Human NOV
21	1393	79.6	263	7	ADN36622 Plasmid p
22	1393	79.6	263	7	ADN36596 Human NOV
23	1306	74.7	451	2	AAK92803 Hepatitis
24	707	40.4	357	8	ADQ76692 Human KIM
25	592	33.8	108	7	ADN36598 Human NOV

26	522.5	29.9	307	2	AAW38334	AAW38334 Rat kidne
27	519	29.7	305	6	ABP70430	Abp70430 Amino aci
28	491.5	28.1	282	6	ABP70431	Abp70431 Amino aci
29	458	26.2	465	8	ADQ76695	Mouse kim
30	432	24.7	81	6	AAO26678	Human KIM
31	405	23.2	345	6	ABP70436	Amino aci
32	401	22.9	305	6	ABP70432	Amino aci
33	401	22.9	305	6	ABP70433	Amino aci
34	401	22.9	345	6	ABP70437	Amino aci
35	367.5	21.0	378	4	AAW39027	Human pol
36	367.5	21.0	378	6	ABP70446	Amino aci
37	367.5	21.0	379	2	AAW25768	Human sec
38	366.5	21.0	378	6	ABP70447	Amino aci
39	321	18.4	183	2	AAW38395	Human sec
40	292	16.7	281	2	AAW01047	Product o
41	292	16.7	281	3	AAW97056	Murine T
42	292	16.7	281	4	AAW50696	Mouse 200
43	292	16.7	281	4	AAW59162	Murine 20
44	292	16.7	281	4	AAW50225	Mouse 200
45	292	16.7	281	4	AAW81511	Murine TH

ALIGNMENTS

RESULT 1
AAW38336
ID AAW38336 standard; protein; 334 AA.

AC AAW38336;

DT 21-MAY-1998 (first entry)

DE Human kidney injury related molecule (KIM).

KW Kidney injury related molecule; KIM; human; renal disease; injury;
nephritis; tissue regeneration; therapy; monoclonal antibody.

OS Homo sapiens.

PN WO9744460-A1.

PD 27-NOV-1997.

PF 23-MAY-1997; 97WO-US009303.

PR 24-MAY-1996; 96US-0018228P.

PR 23-AUG-1996; 96US-0023442P.

XX (BIOJ) BIOGEN INC.

PI Sanicola-Nadel M, Bonventre JV, Hession CA, Ichimura T, Wei H;

PI Cate RL;

XX WPI; 1998-018514/02.

XX DNA encoding kidney injury related molecule - which is upregulated in

XX injured or regenerating tissue, useful to promote growth of new tissue

XX and survival of damaged tissue.

XX Claim 9; Page 46-47; 68pp; English.

XX This protein, designated kidney injury related molecule (KIM), is up-
regulated in injured or regenerating tissue. Its amino acid sequence was
deduced from a clone (see AAW38335) obtained from a human embryonic liver
library. A 572-amino acid (see AAW38335) and a 307-amino acid rat KIM
(see AAW38334) are also claimed. Recombinant KIM polypeptides can be
expressed in prokaryotic and eukaryotic host cells using a claimed
process. Soluble variants fused to a toxin, imageable compound or
radionuclide, and IGG fusion proteins are also claimed. KIM, or an
agonist, can be used to treat renal disease and to promote the growth of new
tissue or the survival of damaged tissue, generally in conditions where
the binding of specific ligand to KIM stimulates cell growth, maintains

cellular differentiation or reduces apoptosis, e.g. in cases of renal failure, nephritis, kidney transplants, toxic or hypoxic injury. A monoclonal antibody specific for KIM can be used to treat renal disease, e.g. where binding of KIM to ligand results in neoplasia, loss of cellular function, susceptibility to apoptosis or promotion of inflammation, deliver imaging agents to KIM expressing cells in vivo or in vitro and measure KIM concentration by immunoassay. Damage/regeneration of renal cells can be determined by measuring KIM, particularly to diagnose or monitor the progress of disease or therapy. KIM-expressing tumour cells can be inhibited by treatment with a fusion protein comprising KIM ligand or MAb with a toxin or radionuclide, and tumour cells that express KIM ligand can be inhibited with similarly tagged KIM or anti-KIM ligand antibody

PI	Baily V, Bonventre J;							
XX	WPI; 2003-156845/15.							
DR	New antibody, antibody derivative or antigen-binding polypeptide that							
XX	inhibits proteolytic release of a soluble kidney injury molecule-1							
PT	polypeptide, useful for treating or preventing renal disease or injury,							
PT	e.g. renal cancer.							
XX								
PS	Disclosure; Page 39-40; 42pp; English.							
XX	The invention relates to a novel antibody, antibody derivative or antigen							
CC	-binding polypeptide that inhibits proteolytic release of a soluble							
CC	kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.							
CC	The antibody, antibody derivative or antigen-binding polypeptide is							
CC	useful for treating or preventing renal disease or injury, e.g. renal							
CC	cancer. The antibody is also useful for inhibiting shedding of the KIM-1							
CC	polypeptide. This sequence represents a protein relating to the KIM-1							
CC	protein of the invention							
XX								
SQ	Sequence 334 AA;							
Query Match		100.0%; Score 1749; DB 6; Length 334;						
Best Local Similarity		100.0%; Pred. No. 7.le-118;						
Matches 334; Conservative		0; Mismatches 0; Indels 0; Gaps 0						
Qy	1	MHPQVTLSLTLLHADS	VAGSVKGGEAGPSVTLPCHYSGAVTSMCWNRGSCSLFTCQG	60				
Dd	1	MHPQVTLSLTLLHADS	VAGSVKGGEAGPSVTLPCHYSGAVTSMCWNRGSCSLFTCQG	60				
Qy	61	IWTNGTHVTRKDTRYKLLGLDLSRRDVSLTIEN	TAVSDSGVYCCEVRHGWFNDMKITV	120				
Dd	61	IWTNGTHVTRKDTRYKLLGLDLSRRDVSLTIEN	TAVSDSGVYCCEVRHGWFNDMKITV	120				
Qy	121	SLEIVPDKVTTTPITVTVPTVVTVTSTVPTTTTVP	TTTTVPTTMMSIPTTTTVPPTMTVS	180				
Dd	121	SLEIVPPKVTTTPITVTVPTVVTVTSTVPTTTTVP	TTTTVPTTMMSIPTTTTVPPTMTVS	180				
Qy	181	TTTTSPVTTTSIPPTTTSVPVTTTVTFVPPMPPLPRONHEPVATSSPOPAEHTPTLQA	240					
Dd	181	TTTTSPVTTTSIPPTTTSVPVTTTVTFVPPMPPLPRQNHEPVATSSPOPAEHTPTLQA	240					
Qy	241	IRREPTSPLYSYTTGDNDVTESDDGLWNNNQTQLFLEHSLLTANTTKGIYAGVCISVL	300					
Dd	241	IRREPTSPLYSYTTGDNDVTESDDGLWNNNQTQLFLEHSLLTANTTKGIYAGVCISVL	300					
Qy	301	VLLALLGVIIAKKYFFKFKEVOOLRPHKSCIHOE	334					
Dd	301	VLLALLGVIIAKKYFFKFKEVOOLRPHKSCIHOE	334					

RESULT 3	
ADQ76696	
ID	ADQ76696 standard; protein; 334 AA.
XX	
AC	ADQ76696;
XX	
DT	16-DEC-2004 (first entry)
XX	
DE	Human kidney injury molecule-1 (KIM-1).
XX	
XX	Human; KIM-1; kidney injury module-1; immunosuppressive; protozoacide;
KW	muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW	antiallergic; antiinflammatory; vasotropic; nephrotropic.
XX	
OS	Homo sapiens.
XX	
XX	
Key	Location/Qualifiers
PH	
FT	Peptide
FT	1..20
FT	/label= Signal_peptide
FT	65..67
FT	/note= "N-glycosylated"
FT	258..260
FT	

FT /note= "N-glycosylated"
 FT 272. .274
 FT /note= "N-glycosylated"
 FT 286. .288
 FT /note= "N-glycosylated"
 FT 290. .311
 FT /note= "Transmembrane domain"
 FT 324. .334
 FT /note= "C-terminal region unique to this splice variant"
 XX
 PN WO2004060041-A2.
 XX
 XX 22-JUL-2004.
 XX
 PF 29-DEC-2003; 2003WO-US041294.
 XX
 PR 30-DEC-2002; 2002US-0436934P.
 XX
 XX (BIOG-) BIOGEN IDEC MA INC.
 XX
 PI Rennert PD;
 XX
 XX WPI; 2004-534277/51.
 XX
 XX Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
 PT Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
 PT activation, or inhibiting production of a subset of antibodies against
 PT one or more antigens.
 XX
 PS Disclosure; SEQ ID NO 1; 68pp; English.
 XX
 CC The present sequence is that of human kidney injury molecule-1 (KIM-1), a
 CC type I cell membrane glycoprotein and member of the T cell immunoglobulin
 CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids or
 CC 359 amino acids ADO76690, depending on splice variation. It has been
 CC discovered that treatment of a mammal with a KIM-1 antagonist alters the
 CC interaction of T cells and other immune system cells, e.g. dendritic
 CC cells, monocytes, macrophages and B cells, and thereby strongly
 CC suppresses an IgG response to an antigen. Such treatment also eliminates
 CC IgG1 production by memory B cells in response to subsequent challenge
 CC with the antigen. Blockage of the binding of KIM-1 to its receptor
 CC reduces secretion of interferon-gamma by immune cells engaged in an
 CC antigen response in the mixed lymphocyte response assay. Based on these
 CC discoveries, the invention provides methods for therapeutically
 CC modulating immune function in autoimmune diseases and other disorders of
 CC the mammalian immune system. A method is claimed for inhibiting
 CC signalling between a T cell and a second cell, e.g. an antigen-presenting
 CC cell, in a mammal. The method involves identifying a mammal with an
 CC immune disease or disorder, or one preparing to receive a tissue graft,
 CC and administering a KIM-1 antagonist, especially a polypeptide
 CC comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a
 CC KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding
 CC fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a
 CC soluble polypeptide, which can include a KIM-1 mucin domain in addition
 CC to the KIM Ig domain, and may include a heterologous moiety such as an Fc
 CC moiety. The antagonist may be conjugated to a polymer such as
 CC polyethylene glycol. The KIM-1 antagonist is used in methods of:
 CC inhibiting activation of a B cell in a mammal; inhibiting disease relapse
 CC in an autoimmune disease; inhibiting epitope spreading in an autoimmune
 CC disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis,
 CC autoimmune haemolytic anaemia, Chagas disease, Grave's disease,
 CC idiopathic thrombocytopenia purpura, Wegener's granulomatosis,
 CC polyarteritis nodosa, rapidly progressive crescentic glomerulonephritis,
 CC graft-versus-host disease, or systemic lupus nephritis) or inflammatory
 CC disease or disorder (e.g. inflammatory bowel diseases such as Crohn's
 CC disease, ulcerative colitis, and ileitis); and inhibiting secretion of
 CC IFN-gamma by lymphocytes.
 XX
 SQ Sequence 334 AA;

Query Match 100.0%; Score 1749; DB 8; Length 334;
 Best Local Similarity 100.0%; Pred. No. 7.1e-118;
 Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPQVILSLILHLADSVAGSVKVGAGPSVTLPCHYSGAVTSMCNRGSCSLFTQNG 60
 Db |||||
 QY 1 MHPQVILSLILHLADSVAGSVKVGAGPSVTLPCHYSGAVTSMCNRGSCSLFTQNG 60
 Db |||||
 QY 61 IWTNGTHYTKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGNFDMKITV 120
 Db |||||
 QY 61 IWTNGTHYTKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGNFDMKITV 120
 Db |||||
 QY 121 SLIEIVPKVTTPIVTTVTVTTRTSTVPTTTTPTTTTPTTTTPTTTTPTTTTPTT 180
 Db |||||
 QY 121 SLIEIVPKVTTPIVTTVTVTTRTSTVPTTTTPTTTTPTTTTPTTTTPTT 180
 Db |||||
 QY 181 TTTSVPTTTSIPPTTSVPVTTTSTVFPPMPPLPRQNHPEVATSPSPQPAETHPTT 240
 Db |||||
 QY 181 TTTSVPTTTSIPPTTSVPVTTTSTVFPPMPPLPRQNHPEVATSPSPQPAETHPTT 240
 Db |||||
 QY 241 IRREPTSSPLSYTTDNDTDTTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
 Db |||||
 QY 241 IRREPTSSPLSYTTDNDTDTTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
 Db |||||
 QY 301 VLLALGVIIIAKKYFFKKEVQQLRPHKSCIHORE 334
 Db |||||
 QY 301 VLLALGVIIIAKKYFFKKEVQQLRPHKSCIHORE 334
 Db |||||
 RESULT 4
 ABR58582
 ID ABR58582 standard; protein; 359 AA.
 XX
 AC ABR58582;
 XX
 DT 09-JUL-2003 (first entry)
 XX
 DE Human cancer related protein SEQ ID NO:239.
 XX
 KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
 KW heart disease; atherosclerosis; endometriosis.
 XX
 OS Homo sapiens.
 XX
 PN WO2003025138-A2.
 XX
 XX 27-MAR-2003.
 PD
 XX 17-SEP-2002; 2002WO-US029560.
 PF
 XX 17-SEP-2001; 2001US-0323469P.
 PR 20-SEP-2001; 2001US-0323887P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 08-FEB-2002; 2002US-0355145P.
 PR 08-FEB-2002; 2002US-0355257P.
 PR 12-APR-2002; 2002US-0372246P.
 XX
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX
 XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
 PI Zlotnik A;
 XX
 XX WPI; 2003-354600/33.
 DR N-PSDB; ACC72709.
 XX
 XX New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 XX
 PS Claim 12; Page 742; 767pp; English.
 XX
 CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in

CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies

XX SQ Sequence 359 AA;
Query Match 96.3%; Score 1684; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.7e-113;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60
Db 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60

Qy 61 IWTNGTHVTYRKDTRYKLLGDLSDRLDVLNTAVSDSGVCCRVHRGWFMNDKITV 120
Db 61 IWTNGTHVTYRKDTRYKLLGDLSDRLDVLNTAVSDSGVCCRVHRGWFMNDKITV 120

Qy 121 SLEIVPPKVTTPITVTTVPTVTSTVPTTTTPTTPTTPTTPTTPTTPTTPTTPTT 180
Db 121 SLEIVPPKVTTPITVTTVPTVTSTVPTTTTPTTPTTPTTPTTPTTPTTPTT 180

Qy 181 TTTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPT 240
Db 181 TTTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPT 240

Qy 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNNTQTLFLEHLLTANTTKGIYAGVCISVL 300
Db 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNNTQTLFLEHLLTANTTKGIYAGVCISVL 300

Qy 301 VLLALLGVIIIAKKYFFKKEVQQL 323
Db 301 VLLALLGVIIIAKKYFFKKEVQQL 323

RESULT 5
ABP70438
ID ABP70438 standard; protein; 359 AA.
XX AC ABP70438;
XX DT 22-APR-2003 (first entry)
XX DE Amino acid sequence of human TIM-1 allele 1.
XX KW T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
XX KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
XX KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
XX KW allergic T cell response; autoimmune disease.
XX OS Homo sapiens.
XX PN WO2003002722-A2.
XX PD 09-JAN-2003.
XX PF 01-JUL-2002; 2002WO-US020890.
XX PR 29-JUN-2001; 2001US-0302344P.
XX PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX WPI; 2003-210268/20.
XX N-ESDB; ABZ68332.
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
XX Mucin domain gene sequences, useful for treating cancer or asthma,
XX allergy, eczema or autoimmune disease.
XX Claim 10; Page 80-81; 94pp; English.
XX The present sequence is a human T cell immunoglobulin domain and mucin
XX domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
XX TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
XX conserved IgV and mucin domains. The locus comprising the TIM family is
XX genetically associated with immune dysfunction, including asthma. The TIM
XX gene family is located within a region of human chromosome 5 that is
XX commonly deleted in malignancies and myelodysplastic syndrome. Variants
XX of TIM-1 and TIM-3 are associated with susceptibility to airway
XX hyperreactivity and allergic T cell responses, and other variants
XX associated with protection against these responses. T cells express TIM
XX proteins, which critically regulate CD4 T cell differentiation. Th1 cells
XX preferentially express TIM-3, while Th2 cells preferentially express TIM-
XX 1. TIM polypeptides and polynucleotides are useful for treating cancer,
XX asthma, allergies, eczema or autoimmune diseases

XX SQ Sequence 359 AA;
Query Match 96.3%; Score 1684; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.7e-113;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60
Db 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCNRGCSLFTCQNG 60

Qy 61 IWTNGTHVTYRKDTRYKLLGDLSDRLDVLNTAVSDSGVCCRVHRGWFMNDKITV 120
Db 61 IWTNGTHVTYRKDTRYKLLGDLSDRLDVLNTAVSDSGVCCRVHRGWFMNDKITV 120

Qy 121 SLEIVPPKVTTPITVTTVPTVTSTVPTTTTPTTPTTPTTPTTPTTPTTPTTPTT 180
Db 121 SLEIVPPKVTTPITVTTVPTVTSTVPTTTTPTTPTTPTTPTTPTTPTTPTT 180

Qy 181 TTTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPT 240
Db 181 TTTTSVPTTTSIPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPTTTSVPT 240

Qy 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNNTQTLFLEHLLTANTTKGIYAGVCISVL 300
Db 241 IRREPTSSPLSYTTDGNDDVTSSDGLWNNNTQTLFLEHLLTANTTKGIYAGVCISVL 300

Qy 301 VLLALLGVIIIAKKYFFKKEVQQL 323
Db 301 VLLALLGVIIIAKKYFFKKEVQQL 323

RESULT 6
ABR48174
ID ABR48174 standard; protein; 359 AA.
XX AC ABR48174;
XX DT 12-JUN-2003 (first entry)
XX DE Human bladder cancer associated protein sequence SEQ ID NO:64.
XX KW Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003003906-A2.

XX	16-JAN-2003.	AAO26680 standard; protein; 359 AA.
PD		
XX		
PF		AAO26680;
XX		
XX	03-JUL-2002; 2002WO-US021338.	
XX		
PR	03-JUL-2001; 2001US-0302814P.	
PR	03-AUG-2001; 2001US-0310099P.	
PR	08-NOV-2001; 2001US-0343705P.	
PR	13-NOV-2001; 2001US-0350666P.	
PR	12-APR-2002; 2002US-0372246P.	
XX		
PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
XX		
PI	Mack DH, Aziz N;	
XX		
DR	WPI; 2003-201532/19.	
DR	N-PSDB; ACC50985.	
XX		
XX		
PT	Detecting a bladder cancer-associated transcript in a cell from a	
PT	patient, comprises contacting a biological sample from the patient with a	
PT	bladder cancer-associated polynucleotide or antibody.	
XX		
XX	Claim 10; Page 252; 307pp; English.	
XX		
CC	The present invention describes a method for detecting a bladder cancer-	
CC	associated transcript in a cell from a patient. The method comprises	
CC	contacting a biological sample from the patient with a polynucleotide	
CC	that selectively hybridizes to a sequence that is 80 % identical to a	
CC	table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059	
CC	encode the human bladder cancer-associated proteins given in ABR48146 to	
CC	ABR48242). Bladder cancer-associated sequences from the present invention	
CC	have cytostatic activities, and can be used in antisense gene therapy and	
CC	in vaccine production. The method can be used for detecting a bladder	
CC	cancer-associated transcript in a cell from a patient. The method is	
CC	useful in diagnosing or treating bladder cancer and in screening for	
CC	compounds that modulate bladder cancer, such as hormones or antibodies.	
CC	The nucleic acid molecules from the present invention may be used in	
CC	various screening and diagnostic methods, and for gene therapy, vaccine	
CC	and/or antisense/inhibition applications	
XX		
SQ	Sequence 359 AA;	
	Query Match 96.3%; Score 1684; DB 6; Length 359;	
	Best Local Similarity 100.0%; Pred. No. 3.7e-113;	
	Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGCSLFTCQNG 60	
DB	1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGCSLFTCQNG 60	
QY	61 IVWTNGTHVYRKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCVHRGFNNDMKITV 120	
DB	61 IVWTNGTHVYRKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCVHRGFNNDMKITV 120	
QY	121 SLEIVPPKVTTPPIVTTVPTVTRTSTVPTTTTPTTPTTMSIPTTTTPTTMTVS 180	
DB	121 SLEIVPPKVTTPPIVTTVPTVTRTSTVPTTTTPTTPTTMSIPTTTTPTTMTVS 180	
QY	181 TTTSVPTTTSIPTTTSVPTTTTSTVFPMPMLPRQNHPEVATSPSPQPAETHPTTLOGA 240	
DB	181 TTTSVPTTTSIPTTTSVPTTTTSTVFPMPMLPRQNHPEVATSPSPQPAETHPTTLOGA 240	
QY	241 IRREPTSSPLSYTTDGNDDVTESDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300	
DB	241 IRREPTSSPLSYTTDGNDDVTESDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300	
QY	301 VLLALLGVIIIAKKYFFKKEVQQL 323	
DB	301 VLLALLGVIIIAKKYFFKKEVQQL 323	
RESULT 7		
AAO26680		

ID	AAO26680 standard; protein; 359 AA.	
XX		
AC	AAO26680;	
XX		
XX		
DT	20-MAR-2003 (first entry)	
XX		
DE	KIM-1 related protein, SEQ ID No 8.	
XX		
KW	Cytostatic; gene therapy; antibody; antigen; antigen-binding;	
KW	proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;	
KW	renal disease; injury; renal cancer; human.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200298920-A1.	
XX		
PD	12-DEC-2002.	
XX		
PF	31-MAY-2002; 2002WO-US017402.	
XX		
PR	01-JUN-2001; 2001US-0295449P.	
PR	04-JUN-2001; 2001US-0295907P.	
XX		
XX	(BIOJ) BIOGEN INC.	
PA	(GEO) GEN HOSPITAL CORP.	
XX		
PI	Bailly V, Bonventre J;	
XX		
DR	WPI; 2003-156845/15.	
XX		
PT	New antibody, antibody derivative or antigen-binding polypeptide that	
PT	inhibits proteolytic release of a soluble kidney injury molecule-1	
PT	polypeptide, useful for treating or preventing renal disease or injury,	
PT	e.g. renal cancer.	
XX		
PS	Disclosure; Page 40-41; 42pp; English.	
XX		
CC	The invention relates to a novel antibody, antibody derivative or antigen	
CC	-binding polypeptide that inhibits proteolytic release of a soluble	
CC	kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.	
CC	The antibody, antibody derivative or antigen-binding polypeptide is	
CC	useful for treating or preventing renal disease or injury, e.g. renal	
CC	cancer. The antibody is also useful for inhibiting shedding of the KIM-1	
CC	polypeptide. This sequence represents a protein relating to the KIM-1	
CC	protein of the invention	
XX		
SQ	Sequence 359 AA;	
	Query Match 96.3%; Score 1684; DB 6; Length 359;	
	Best Local Similarity 100.0%; Pred. No. 3.7e-113;	
	Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGCSLFTCQNG 60	
DB	1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGCSLFTCQNG 60	
QY	61 IVWTNGTHVYRKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCVHRGFNNDMKITV 120	
DB	61 IVWTNGTHVYRKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCVHRGFNNDMKITV 120	
QY	121 SLEIVPPKVTTPPIVTTVPTVTRTSTVPTTTTPTTPTTMSIPTTTTPTTMTVS 180	
DB	121 SLEIVPPKVTTPPIVTTVPTVTRTSTVPTTTTPTTPTTMSIPTTTTPTTMTVS 180	
QY	181 TTTSVPTTTSIPTTTSVPTTTTSTVFPMPMLPRQNHPEVATSPSPQPAETHPTTLOGA 240	
DB	181 TTTSVPTTTSIPTTTSVPTTTTSTVFPMPMLPRQNHPEVATSPSPQPAETHPTTLOGA 240	
QY	241 IRREPTSSPLSYTTDGNDDVTESDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300	
DB	241 IRREPTSSPLSYTTDGNDDVTESDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300	
QY	301 VLLALLGVIIIAKKYFFKKEVQQL 323	

Db 301 VLLALLGVIIIAKYPFKKEVQQL 323
RESULT 8
ADE36594
ID ADE36594 standard; protein; 359 AA.
XX ADE36594;
AC ADE36594;
XX 29-JAN-2004 (first entry)
XX Human NOV1b protein SEQ ID NO:4.
DE human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
XX renal cancer; inflammation; tissue typing.
KW Homo sapiens.
XX WO2003080856-A2.
PN 02-OCT-2003.
XX 19-MAR-2003; 2003WO-US008490.
XX 19-MAR-2002; 2002US-0365491P.
PR 13-SEP-2002; 2002US-0410618P.
XX (CURA-) CURAGEN CORP.
PA Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
XX Ooi CE, Anderson DW, Guo X, Giot L, Starling G;
PI WPI; 2003-876927/81.
DR N-PSDB; ADE36593.
XX New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
XX Claim 1; SEQ ID NO 4; 239pp; English.
XX The present invention describes an isolated human NOVX polypeptide, where
CC X is la to ld or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to it, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOV1b from the present invention.
XX Sequence 359 AA;
SQ
Query Match 96.3%; Score 1684; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 3.7e-113; Mismatches 0; Gaps 0;
Matches 323; Conservative 0; Indels 0;
QY 1 MHPQVVILSLIHLADSVAGSVKVGEGAPSVTLPLCHYSGAVTSCMNRGCSLFTCCQNG 60
Db 1 MHPQVVILSLIHLADSVAGSVKVGEGAPSVTLPLCHYSGAVTSCMNRGCSLFTCCQNG 60
QY 61 IWTNGTHVTRKDYRKLGLDLSRRDVSLSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Db 61 IWTNGTHVTRKDYRKLGLDLSRRDVSLSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPPIVTTVPTVTSTVPTTTPVTTPVTMSIPTTTVPTMTVS 180
Db 121 SLEIVPPKVTTPPIVTTVPTVTSTVPTTTPVTTPVTMSIPTTTVPTMTVS 180
QY 181 TTTSVPTTTSIPTTTSVPVPTTTSVTFVPPMPLPRQNHEPVATSPSSPQPAETHPTTLOGA 240
Db 181 TTTSVPTTTSIPTTTSVPVPTTTSVTFVPPMPLPRQNHEPVATSPSSPQPAETHPTTLOGA 240

Db 181 TTTSVPTTTSIPTTTSVPVPTTTSVTFVPPMPLPRQNHEPVATSPSSPQPAETHPTTLOGA 240
QY 241 IRREPTSSPLSYTTDGNDFVTTESDGLWNNNOTQLFLEHSLLTANTTKGIYAGVCISVL 300
Db 241 IRREPTSSPLSYTTDGNDFVTTESDGLWNNNOTQLFLEHSLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIIAKYPFKKEVQQL 323
Db 301 VLLALLGVIIIAKYPFKKEVQQL 323
RESULT 9
ADN38984
ID ADN38984 standard; protein; 359 AA.
XX ADN38984;
AC ADN38984;
XX 17-JUN-2004 (first entry)
XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:302.
DE Human; differential expression; cancer; angiogenic disorder;
XX fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularistation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine.
XX Homo sapiens.
OS WO2003042661-A2.
PN 22-MAY-2003.
XX 13-NOV-2002; 2002WO-US036810.
XX 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN38983.
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX Claim 12; SEQ ID NO 302; 1385pp; English.
PS The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods

RESULT 10	
ADQ76690	
ID	ADQ76690 standard; protein; 359 AA.
XX	
AC	ADQ76690;
XX	
DT	16-DEC-2004 (first entry)
XX	
DE	Human kidney injury molecule-1 (KIM-1).
XX	
KW	Human; KIM-1; kidney injury module-1; immunosuppressive; protozoacide;
KW	muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW	antiallergic; antiinflammatory; vasotropic; nephropic.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	1..20
FT	/label= Signal_peptide
FT	65..67
FT	/note= "N-glycosylated"
FT	258..260
FT	/note= "N-glycosylated"
FT	272..274
FT	/note= "N-glycosylated"
FT	286..288
FT	/note= "N-glycosylated"
FT	

XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
XX
PS Claim 10; Page 84-85; 94pp; English.
XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IgV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
XX Sequence 359 AA;
Query Match 95.5%; Score 1671; DB 6; Length 359;
Best Local Similarity 99.1%; Pred. No. 3.2e-112;
Matches 320; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGSCSLFTCQNG 60
Db 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGSCSLFTCQNG 60
QY 61 IVWTNGTHVTRKOTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Db 61 IVWTNGTHVTRKOTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPPIVTTVPTVTVTSTVPTTTTPTTTTPTTTTPTTTTPTTTTPTTVS 180
Db 121 SLEIVPPKVTTPPIVTTVPTVTVTSTVPTTTTPTTTTPTTTTPTTTTPTTTTPTTVS 180
QY 181 TTTSVPTTTSIPITTSVPVTTVSTVTPMPPLPRQNHPEVATSPSPQPAETHPTTLOGA 240
Db 181 TTTSVPTTTSIPITTSVPVTTVSTVTPMPPLPRQNHPEVATSPSPQPAETHPTTLOGT 240
QY 241 IRREPTSSPLSYTTDNDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
Db 241 IRREPTSSPLSYTTDNDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
QY 301 VLLALLGVIIAKKYFFKKEVQQL 323
Db 301 VLLALLGVIIAKKYFFKKEVQQL 323
RESULT 13
ABP70440
ID ABP70440 standard; protein; 365 AA.
XX
AC ABP70440;
XX
DT 22-APR-2003 (first entry)
XX
DE Amino acid sequence of human TIM-1 allele 3.
XX
DE T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 208
FT /note= "encoded by RCG"

XX WO2003002722-A2.
XX
PD 09-JAN-2003.
XX
PP 01-JUL-2002; 2002WO-US020890.
XX
PR 29-JUN-2001; 2001US-0302344P.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX WPI; 2003-210268/20.
DR N-PSDB; ABZ68334.
XX
PT New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
XX
PS Claim 10; Page 83-84; 94pp; English.
XX
CC The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IgV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
XX Sequence 365 AA;
Query Match 95.3%; Score 1666; DB 6; Length 365;
Best Local Similarity 97.9%; Pred. No. 7.5e-112;
Matches 322; Conservative 0; Mismatches 1; Indels 6; Gaps 1;
QY 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGSCSLFTCQNG 60
Db 1 MHPQVILSLIHLADSVAGSVKVGEGAPSVTLPCHYSGAVTSMCWNRGSCSLFTCQNG 60
QY 61 IVWTNGTHVTRKOTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
Db 61 IVWTNGTHVTRKOTRYKLLGLDLSRDDVSLTIENTAVSDSGVYCCRVHRGFNDMKITV 120
QY 121 SLEIVPPKVTTPPIVTTVPTVTVTSTVPTTTTPTTTTPTTTTPTTTTPTTTTPTTV 174
Db 121 SLEIVPPKVTTPPIVTTVPTVTVTSTVPTTTTPTTTTPTTTTPTTTTPTTTTPTTV 180
QY 175 TTMTVSTTTSVPTTTSIPITTSVPVTTVSTVTPMPPLPRQNHPEVATSPSPQPAETHP 234
Db 181 TTMTVSTTTSVPTTTSIPITTSVPVTTVSTVTPMPPLPRQNHPEVATSPSPQPAETHP 240
QY 235 TTLOGAIRREPTSSPLSYTTDNDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAG 294
Db 241 TTLOGAIRREPTSSPLSYTTDNDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAG 300
QY 295 VCISVLVLLALLGVIIAKKYFFKKEVQQL 323
Db 301 VCISVLVLLALLGVIIAKKYFFKKEVQQL 329
RESULT 14
ABP70442
ID ABP70442 standard; protein; 364 AA.
XX
AC ABP70442;

XX 22-APR-2003 (first entry)
DT Amino acid sequence of human TIM-1 allele 5.
XX
XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
XX WO2003002722-A2.
PN
XX
XX 09-JAN-2003.
XX
XX 01-JUL-2002; 2002WO-US020890.
PF
XX 29-JUN-2001; 2001US-0302344P.
PR
XX (STRD) UNIV LELAND STANFORD JUNIOR.
PA
XX
XX McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
PI
XX
XX MPI; 2003-210268/20.
DR
XX N-PSDB; ABZ68336.
DR
XX
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
PT
XX
XX Claim 10; Page 86-87; 94pp; English.
PS
XX
XX The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IgV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
XX Sequence 364 AA;
PS
XX
XX Query Match 94.7%; Score 1655.5; DB 6; Length 364;
Best Local Similarity 97.9%; Pred. No. 4.3e-111;
Matches 322; Conservative 0; Mismatches 0; Indels 7; Gaps 2;
QY 1 MHPQVILSLIHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSCWNRGSCSLFTCCQNG 60
DB 1 MHPQVILSLIHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSCWNRGSCSLFTCCQNG 60
QY 61 IVWTNGTHVTRKDTYKLLGLDLSRRDVSLSLTIENTAVSDSGVYCCRVHRGWFNDMKITV 120
DB 61 IVWTNGTHVTRKDTYKLLGLDLSRRDVSLSLTIENTAVSDSGVYCCRVHRGWFNDMKITV 120
QY 121 SLEIIVPPKVTTPITVTTVTVTVTSTVTP-----TTTTVTVTVTTPITTTTTP 174
DB 121 SLEIIVPPKVTTPITVTTVTVTVTSTVTP-----TTTTVTVTVTTPITTTTTP 180
QY 175 TMTVSTTTSPTTTSIPTTSVPVTTTSTVTPVPMPLPRQNHPEVATSPSPQAEATHP 234
DB 181 TMTVSTTTSPTTTSIPTTSVPVTTTSTVTPVPMPLPRQNHPEVATSPSPQAEATHP 239
QY 235 TTLOGAIRREPTSPPLSYTTDGNVTTESSDGLWNNNQTLFLEHSLLTANTTKGIYAG 294
DB 240 TTLOGAIRREPTSPPLSYTTDGNVTTESSDGLWNNNQTLFLEHSLLTANTTKGIYAG 299

QY 295 VCISVLVLLALLGVIIIAKKYFFKKEVQQL 323
DB 300 VCISVLVLLALLGVIIIAKKYFFKKEVQQL 328
RESULT 15
ABP70443
ID ABP70443 standard; protein; 364 AA.
XX
XX AC ABP70443;
XX
XX 22-APR-2003 (first entry)
XX
XX Amino acid sequence of human TIM-1 allele 6.
XX
XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW allergic T cell response; autoimmune disease.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 123 /note= "encoded by GAG"
FT Misc-difference 161 /note= "encoded by GACT"
FT Misc-difference 201 /note= "encoded by ACAAGT"
FT
XX WO2003002722-A2.
XX
XX 09-JAN-2003.
XX
XX 01-JUL-2002; 2002WO-US020890.
XX
XX 29-JUN-2001; 2001US-0302344P.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX
XX McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
PI
XX
XX MPI; 2003-210268/20.
DR
XX N-PSDB; ABZ68336.
DR
XX
XX New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT Mucin domain gene sequences, useful for treating cancer or asthma,
PT allergy, eczema or autoimmune disease.
PT
XX
XX Claim 10; Page 87-88; 94pp; English.
PS
XX
XX The present sequence is a human T cell immunoglobulin domain and mucin
CC domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC conserved IgV and mucin domains. The locus comprising the TIM family is
CC genetically associated with immune dysfunction, including asthma. The TIM
CC gene family is located within a region of human chromosome 5 that is
CC commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC of TIM-1 and TIM-3 are associated with susceptibility to airway
CC hyperreactivity and allergic T cell responses, and other variants
CC associated with protection against these responses. T cells express TIM
CC proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC 1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC asthma, allergies, eczema or autoimmune diseases
XX
XX Sequence 364 AA;
PS
XX
XX Query Match 94.0%; Score 1643.5; DB 6; Length 364;
Best Local Similarity 97.3%; Pred. No. 3.1e-110;
Matches 320; Conservative 1; Mismatches 1; Indels 7; Gaps 2;

Qy	1	MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCWNRGSCSLFTCQNG	60
Db	1	MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSMCWNRGSCSLFTCQNG	60
Qy	61	IWVTNGTHVYRKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGWENDMKITV	120
Db	61	IWVTNGTHVYRKDTRYKLLGDLRRDVSLLTIENTAVSDSGVYCCRVHRGWENDMKITV	120
Qy	121	SLEIVPPKVTTPDIVTTVPTVTTVTRSTTVPT-----TTTVPTTTVPTTMSIPTTTTVP	174
Db	121	SLGIVPPKVTTPDIVTTVPTVTTVTRSTTVPTTTTVPTTVPMTTTPVPTTMSIPTTTTVP	180
Qy	175	TTMTVSTTTSVPPTTSIPTTTSVPVTTVSTFVPPMPLPRONHEPVATSPSSQPAETHP	234
Db	181	TTMTVSTTTSVPPTTSIP-TTISVPVTTVSTFVPPMPLPRONHEPVATSPSSQPAETHP	239
Qy	235	TTLOGAIRREPTSSPLYSVTTDGNDTVTESSDGLWNNNOTQLFLEHSLLTANTTKGIYAG	294
Db	240	TTLOGAIRREPTSSPLYSVTTDGDVTVTESSDGLWNNNOTQLFLEHSLLTANTTKGIYAG	299
Qy	295	VCISVLVLLALLGVIIAKKYFFKKEVQQL	323
Db	300	VCISVLVLLALLGVIIAKKYFFKKEVQQL	328

Search completed: June 29, 2005, 09:09:08
Job time : 238.54 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 133.862 Seconds
(without alignments)
1277.688 Million cell updates/sec

Title: US-10-718-321-7
Perfect score: 1749
Sequence: 1 MHPQVILSLILHLADSVAG.....FPKKEVQQLRPHKSCIHQRE 334

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_eprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1684	96.3	359	2	O43656 homo sapien
2	1661.5	95.0	364	2	O96D42 homo sapien
3	1309	74.8	451	2	O95144 cercopithec
4	1303.5	74.5	460	2	O18984 cercopithec
5	1288	73.6	469	2	O7JJ48 cercopithec
6	1288	73.6	474	2	O46597 cercopithec
7	1286	73.5	473	2	O7JJ47 cercopithec
8	1286	73.5	478	2	O46598 cercopithec
9	522.5	29.9	307	2	O54947 rattus norv
10	519	29.7	305	2	O8VIM2 mus musculu
11	491.5	28.1	282	2	O8VIM1 mus musculu
12	486.5	27.8	282	2	O7TFU2 mus musculu
13	401	22.9	305	2	O8VBW0 mus musculu
14	401	22.9	362	2	O8CIC7 mus musculu
15	399	22.8	343	2	O6U7R3 mus musculu
16	398	22.8	343	2	O6U7R4 mus musculu
17	392	22.4	305	2	O8R183 mus musculu
18	385	22.0	438	2	O6INH8 xenopus lae
19	367.5	21.0	378	2	O96H15 homo sapien
20	283	16.2	281	2	O8VIM0 mus musculu
21	269	15.4	192	2	O7TQ03 mus musculu
22	267	15.3	301	2	O96K94 homo sapien
23	266	15.2	301	2	O8TDQ0 homo sapien
24	254	14.5	224	2	O6G6V2 xenopus lae
25	250	14.3	582	2	O8IR74 drosophila
26	246.5	14.1	662	1	MUC1_XENLA
27	246	14.1	538	2	O76H84 streptococ
28	244	14.0	400	1	MUA1_XENLA
29	238	13.6	483	2	O9W4M2 drosophila
30	235.5	13.5	142	2	O8WW60 homo sapien
31	235.5	13.5	468	2	O55279 avian adeno

ALIGNMENTS

RESULT 1

O43656 PRELIMINARY; PRT; 359 AA.

AC O43656:

DT 01-JUN-1998 (Tremblrel. 06, Created)

DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

DE Hepatitis A virus cellular receptor 1.

GN Name=hAVcr-1;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RC MEDLINE=98325180; PubMed=9658108;

RA Feigelstock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;

RT "The human homolog of HAVcr-1 codes for a hepatitis A virus cellular

RT receptor.";

RL J. Virol. 72:6621-6628(1998).

DR EMBL: AF043724; AAC39862.1; -.

DR GO: GO:0004872; Fireceptor activity; IEA.

DR InterPro: IPR003599; IG.

DR InterPro: IPR007110; IG-like.

DR SMART: SM00409; IG_MHC.

DR PROSITE: PS50835; IG_LIKE; 1.

DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.

DR RECEPTOR.

DR SEQUENCE 359 AA; 38704 MW; C207FEC562DC62CA CRC64;

Query Match 96.3%; Score 1684; DB 2; Length 359;

Best Local Similarity 100.0%; Pred. No. 5.1e-101; Mismatches 0; Indels 0; Gaps 0;

Matches 323; Conservative 0;

QY 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCNRGSCSLFTCQNG 60

DB 1 MHPQVILSLILHLADSVAGSVKVGGEAGPSVTLPCYSGAVTSMCNRGSCSLFTCQNG 60

QY 61 IWTNGTHTYRKDTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCVHRGFNDMKITV 120

DB 61 IWTNGTHTYRKDTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCVHRGFNDMKITV 120

QY 121 SLEIIVPPKVTTPITVPTVTVTSTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

DB 121 SLEIIVPPKVTTPITVPTVTVTSTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180

QY 181 TTTSVPTTTSIPTTTTSVPVTTTSTVFPMPPLPRQNHPEVATSPSPQPAETHPTTLOQA 240

DB 181 TTTSVPTTTSIPTTTTSVPVTTTSTVFPMPPLPRQNHPEVATSPSPQPAETHPTTLOQA 240

QY 241 TRREPTSPPLYSYTTDGDNDVTTESSDGLNWNQTLFLEHLLTANTTKGIAGVCISVL 300

DB 241 TRREPTSPPLYSYTTDGDNDVTTESSDGLNWNQTLFLEHLLTANTTKGIAGVCISVL 300

```

Db      241  IRREPTSSPLSYTTDGNDDTVTSSDGLWNNNOTQLFLEHSLLTANTTKGIYAGVCISVL 300
QY      301  VLLALLGVIIAKKYFFKKEVQQL 323
Db      301  VLLALLGVIIAKKYFFKKEVQQL 323

RESULT 2
Q96D42  ID Q96D42      PRELIMINARY;      PRT;      364 AA.
AC      Q96D42;
DT      01-DEC-2001 (TRENBLrel. 19, Created)
DT      01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT      25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE      HAVCR1 protein.
GN      Name=HAVCR1;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA      Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA      Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA      Jones S.J., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      Strausberg R.;
RL      Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RA      Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL      Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR      EMBL; BC013325; AAH13325.1; -.
DR      EMBL; CR457114; CAG33395.1; -.
DR      Genew; HGNC:17866; HAVCR1.
DR      InterPro; IPR003599; IG.
DR      InterPro; IPR007110; IG-like.
DR      InterPro; IPR003006; IG_MHC.
DR      SMART; SM00409; IG; 1.
DR      PROSITE; PS50835; IG LIKE; 1.
DR      PROSITE; PS00290; IG_MHC; UNKNOWN 1.
SQ      SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;

Query Match      95.0%; Score 1661.5; DB 2; Length 364;
Best Local Similarity 98.2%; Pred. No. 1.5e-99;
Matches 322; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY      1  MHPQVILSLIHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSCMWRGCSLFTCCNG 60
Db      1  MHPQVILSLIHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSCMWRGCSLFTCCNG 60
QY      61  IVWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNMDKITV 120

us-10-718-321-7.rup
Db      61  IVWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNMDKITV 120
QY      121  SLEIVPPKVTTPPIVTTVPTTTRTSTTVPTT-----TTVPTTTPPTTMSIPTTTTVP 175
Db      121  SLEIVPPKVTTPPIVTTVPTTTRTSTTVPTT-----TTVPTTTPPTTMSIPTTTTVP 180
QY      176  TMTVSTTTTSVPTTTSIPPTTTSVPVTTTSTVTPVPMPLPRONHEPVATSPSSQPAPETHT 235
Db      181  TMTVSTTTTSVPTTTSIPPTTTSVPVTTTSTVTPVPMPLPRONHEPVATSPSSQPAPETHT 240
QY      236  TLQGAIRREPTSSPLSYTTDGNDDTVTSSDGLWNNNOTQLFLEHSLLTANTTKGIYAGV 295
Db      241  TLQGAIRREPTSSPLSYTTDGNDDTVTSSDGLWNNNOTQLFLEHSLLTANTTKGIYAGV 300
QY      296  CISVLVLLALLGVIIAKKYFFKKEVQQL 323
Db      301  CISVLVLLALLGVIIAKKYFFKKEVQQL 328

RESULT 3
Q95144  ID Q95144      PRELIMINARY;      PRT;      451 AA.
AC      Q95144;
DT      01-FEB-1997 (TRENBLrel. 02, Created)
DT      01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT      01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE      HAVcr-1 protein precursor.
OS      Cercopithecus aethiops (Green monkey) (Grivet).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC      Cercopithecinae; Cercopithecus.
OX      NCBI_TaxID=9534;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Kidney;
RX      MEDLINE=97015129; PubMed=8861957;
RA      Kaplan G., Totsuka A., Thompson P., Akatsuka T., Moritsugu Y.,
RA      Feinstein S.M.;
RA      "Identification of a surface glycoprotein on African green monkey
RT      kidney cells as a receptor for hepatitis A virus.";
RL      EMBO J. 15:4282-4296(1996).
DR      EMBL; X98252; CAA66906.1; -.
DR      PIR; S71754; S71754.
DR      InterPro; IPR003599; IG.
DR      InterPro; IPR007110; IG-like.
DR      InterPro; IPR003006; IG_MHC.
DR      SMART; SM00409; IG; 1.
DR      PROSITE; PS50835; IG LIKE; 1.
DR      PROSITE; PS00290; IG_MHC; UNKNOWN 1.
KW      Signal.
FT      SIGNAL.
SQ      SEQUENCE 451 AA; 48774 MW; 5D395C5455AA4332 CRC64;

Query Match      74.8%; Score 1309; DB 2; Length 451;
Best Local Similarity 63.3%; Pred. No. 1e-76;
Matches 255; Conservative 36; Mismatches 32; Indels 80; Gaps 3;

QY      1  MHPQVILSLIHLADSVAGSVKVGGEAGPSVTLPCHYSGAVTSCMWRGCSLFTCCNG 60
Db      1  MHPQVILSLIHLADSVAGSVKVGGEAGPSVTLPCRYNGAITSCMWRGCSVFCSDPG 60
QY      61  IVWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNMDKITV 120
Db      61  IVWTNGTHVTRKOTRYKLLGDLSDRDVSLTIENTAVSDSGVYCCRVHRGFNMDKITV 120
QY      121  SLEIVPPKV-----TTTPIVTTVP 139
Db      121  SLKIGPPRVTPIVRTVRTSTTVPTTTTLPPTTTTLPMTTLPMTTLP 180
QY      140  TTTVTRSTTVPTTTTVP-----TTTVPPTTMSIPTTTTPTTMTVSTT 182
Db      181  TTTVPTTTLPTTTTLPMTTLPMTTLPMTTLPMTTLPMTTLPMTTLP 240

```

```
QY 183 T-----SVPTTSPITTSVPTVTTTSTVTFVPPMPLPRQNHVPV 220
Db 241 TTTTPTTTLPTMTLPTTTTPTTMTLPTTTTPTTTLPTTTTPTTMTVSTFVPPPTPLPMQNHEPV 300

QY 221 ATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNDFVTESSDGLMNNQTLFLEH 280
Db 301 ATSPSSQPAETHPTVLLGATRTQPTSSPLSYTTDGSDFVTESSDGLMNNQTLSPHE 360

QY 281 SLTANTTKGIYAGVCISVLVLLALLGVIIAKKYFFKKEVQOL 323
Db 361 SPQWNTTEGIYAGVCISVLVLLAVLGVIYAKKYFFKKEIQOL 403

RESULT 4
O18984 PRELIMINARY; PRT; 460 AA.
AC O18984;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-MAR-2004 (T-EMBLrel. 26, Last sequence update)
DE Hepatitis A virus receptor.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97368416; PubMed=9225030;
RA Ashida M., Hamada C.;
RT "Molecular cloning of the hepatitis A virus receptor from a simian
RT cell line.";
RL J. Gen. Virol. 78:1565-1569(1997).
DR EMBL; D88585; BAA21556.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 460 AA; 49687 MW; 1D1A0C1D832139EA CRC64;

Query Match 74.5%; Score 1303.5; DB 2; Length 460;
Best Local Similarity 61.9%; Pred. No. 2.4e-76;
Matches 255; Conservative 35; Mismatches 33; Indels 89; Gaps 3;

QY 1 MHPQVILSLILHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCNRGSCSLFTCQNG 60
Db 1 MHLQVILSLILHLADFVADSVNVDGVLGSLITLPCRYNGAITSMCNMGTCVFSFCDG 60

QY 61 IWTNGTHVTRKDYKLLGLDLSRDDVSLTENTAVSDSGVCCVCRVHRGFNDMKITV 120
Db 61 IWTNGTHVTRKDYKLLGLNLSRDDVSLTANTAVSDSGVCCVCRVHRGFNDMKITI 120

QY 121 SLEIVPPKV-----TTPTVTVPTVTVTRSTVPTTTTPTTTLPTTTLPTTTLPTT 157
Db 121 SUKIGPRVTIPIVTRVTRSTVPTTTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 180

QY 158 ---TTTVP-----TTTVP-----TTTVP-----TTTVP-----TTTVP 163
Db 181 LPTTTTVPMTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 240

QY 164 -----TMSIPTTTTPTMTVSTVSTVPTTTSIPTTTSVPTTTSVTFVPPMP 211
Db 241 PTTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTT 300

QY 212 LPRQNHVPATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNDFVTESSDGLMNN 271
Db 301 LPMQNHVPATSPSSQPAETHPTVLLGATRTQPTSSPLSYTTDGSDFVTESSDGLMNN 360
```

```
QY 272 NOTQLFLEHSLLTANTTKGIYAGVCISVLVLLALLGVIIAKKYFFKKEVQOL 323
Db 361 NOTQLSPEHSPQWNTTEGIYAGVCISVLVLLAVLGVIYAKKYFFKKEIQOL 412

RESULT 5
Q7JJ48 PRELIMINARY; PRT; 469 AA.
AC Q7JJ48;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Hepatitis A virus cellular receptor 1 short form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043446; AAC39771.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 469 AA; 50561 MW; 27E386F47948F528 CRC64;

Query Match 73.6%; Score 1288; DB 2; Length 469;
Best Local Similarity 59.9%; Pred. No. 2.5e-75;
Matches 252; Conservative 39; Mismatches 32; Indels 98; Gaps 3;

QY 1 MHPQVILSLILHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCNRGSCSLFTCQNG 60
Db 1 MHLQVILSLILHLADSVADSVNVDGVLGSLITLPCRYNGAITSMCNMGTCVFSFCDG 60

QY 61 IWTNGTHVTRKDYKLLGLDLSRDDVSLTENTAVSDSGVCCVCRVHRGFNDMKITV 120
Db 61 IWTNGTHVTRKDYKLLGLNLSRDDVSLTANTAVSDSGVCCVCRVHRGFNDMKITI 120

QY 121 SLEIVPPKV-----TTTVP-----TTTVP-----TTTVP-----TTTVP 130
Db 121 SUKIGPRVTIPIVTRVTRSTVPTTTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 180

QY 131 -----TTTVP-----TTTVP-----TTTVP-----TTTVP-----TT 147
Db 181 LPTTTTVPMTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPT 240

QY 148 TTVPTTTTVP-----TMSIPTTTTPTMTVSTVSTVPTTTSIPTTTSVPTT 202
Db 241 TTVPTTTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTTTLPTT 300

QY 203 VSTFVPPMPLPRQNHVPATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNDFV 262
Db 301 VSTFVPPMPLPRQNHVPATSPSSQPAETHPTVLLGATRTQPTSSPLSYTTDGNDFV 360

QY 263 ESSDGLMNNQTLFLEHSLLTANTTKGIYAGVCISVLVLLALLGVIIAKKYFFKKEVQ 322
Db 361 ESSDGLMNNQTLSPHEHSPQWNTTEGIYAGVCISVLVLLAVLGVIYAKKYFFKKEIQ 420

QY 323 L 323
Db 421 L 421
```

RESULT 6
 O46597 PRELIMINARY; PRT; 474 AA.
 AC O46597; 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hepatitis A virus cellular receptor 1 long form.
 GN Name=HAVcr-1;
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98285791; PubMed=9621093;
 RA Feigelshtock D., Thompson P., Mattos P., Kaplan G.G.;
 RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
 RT green monkey kidney cells result in antigenic variants that do not
 RT react with protective monoclonal antibody 190/4.";
 RL J. Virol. 72:6218-6222(1998).
 DR EMBL; AF043448; AAC39773.1; -;
 GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Receptor.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 SQ SEQUENCE 474 AA; 51088 MW; 472DF987CA220524 CRC64;
 Query Match 73.6%; Score 1288; DB 2; Length 474;
 Best Local Similarity 59.9%; Pred. No. 2.5e-75;
 Matches 252; Conservative 39; Mismatches 32; Indels 98; Gaps 3;
 QY 1 MHPQVILSLHLADSVAGSVKVGAGPSVTLPCYSGAVTSMCWNRGCSLFTCONG 60
 DB 6 MHLQVILSLHLADSVAGSVKVGAGPSVTLPCYSGAVTSMCWNRGCSLFTCONG 65
 QY 61 IVWNGTHVYRKDTRYKLLGDLSSRDVSLTIENTAVSDSGVYCCRVHGRGFWNDMKITV 120
 DB 66 IVWNGTHVYRKDTRYKLLGDLSSRDVSLTIENTAVSDSGVYCCRVHGRGFWNDMKITV 125
 QY 121 SLEIVPKVT----- 130
 DB 126 SLKIGPRVTPIVTRVTRSTVPTTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTT 185
 QY 131 -----TTPIVTVPTVTRTS 147
 DB 186 LPTTTTVPMTTTTLPTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTT 245
 QY 148 TTPPTTTVP-TTTPVT-----TMSIPTTTVPTMTVSTTSTVPTTTSIPTTSVPVTTT 202
 DB 246 TTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTT 305
 QY 203 VSTFVPMPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVT 262
 DB 306 VSTFVPMPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVT 365
 QY 263 ESSDGLNNNTQQLFLEHSLTANTTKGIYAGVCISVLVLLGLVIAKKYFFKVEQQ 322
 DB 366 ESSDGLNNNTQQLFLEHSLTANTTKGIYAGVCISVLVLLGLVIAKKYFFKVEQQ 425
 QY 323 L 323
 DB 426 L 426
 RESULT 7
 O46598 PRELIMINARY; PRT; 478 AA.
 AC O46598; 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hepatitis A virus cellular receptor 1 short form.
 GN Name=HAVcr-1;
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98285791; PubMed=9621093;
 RA Feigelshtock D., Thompson P., Mattos P., Kaplan G.G.;
 RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
 RT green monkey kidney cells result in antigenic variants that do not
 RT react with protective monoclonal antibody 190/4.";
 RL J. Virol. 72:6218-6222(1998).
 DR EMBL; AF043448; AAC39773.1; -;
 GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Receptor.
 DR PROSITE; PS00835; IG LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 SQ SEQUENCE 473 AA; 50973 MW; CD15EF5EE79C8013 CRC64;
 Query Match 73.5%; Score 1286; DB 2; Length 473;
 Best Local Similarity 58.9%; Pred. No. 3.3e-75;
 Matches 251; Conservative 38; Mismatches 33; Indels 104; Gaps 2;
 QY 1 MHPQVILSLHLADSVAGSVKVGAGPSVTLPCYSGAVTSMCWNRGCSLFTCONG 60
 DB 1 MHLQVILSLHLADSVAGSVKVGAGPSVTLPCYSGAVTSMCWNRGCSLFTCONG 60
 QY 61 IVWNGTHVYRKDTRYKLLGDLSSRDVSLTIENTAVSDSGVYCCRVHGRGFWNDMKITV 120
 DB 61 IVWNGTHVYRKDTRYKLLGDLSSRDVSLTIENTAVSDSGVYCCRVHGRGFWNDMKITV 120
 QY 121 SLEIVPKVT----- 129
 DB 121 SLKIGPRVTPIVTRVTRSTVPTTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTT 180
 QY 130 -----TTPIVTVPTVTRTS 137
 DB 181 LPTTTTVPMTTTTLPTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTT 240
 QY 138 VPVTVTVRSTVPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTT 197
 DB 241 LPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTTTLPTTT 299
 QY 198 PVTTTSTVTPVPMPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTD 257
 DB 300 PTTTMTSTVTPVPMPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTD 359
 QY 258 NDTVTSSDGLNNNTQQLFLEHSLTANTTKGIYAGVCISVLVLLGLVIAKKYFFK 317
 DB 360 SDIVTSSDGLNNNTQQLFLEHSLTANTTKGIYAGVCISVLVLLGLVIAKKYFFK 419
 QY 318 KEVQQL 323
 DB 420 KEIQQQL 425
 RESULT 8
 O46598
 ID O46598
 AC O46598;

RESULT 7

DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hepatitis A virus cellular receptor 1 long form.
 GN Name=HAVcr-1;
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98285791; PubMed=9621093;
 RA Feigelscock D., Thompson P., Mattoo P., Kaplan G.G.;
 RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
 RT green monkey kidney cells result in antigenic variants that do not
 RT react with protective monoclonal antibody 190/4.";
 RL J. Virol. 72:6218-6222(1998).
 DR EMBL; AF043449; AAC39774.1; -;
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR SMART; SM00409; IG_1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
 KW Receptor.
 SQ SEQUENCE 478 AA; 51501 MW; 3A919655C752FF60 CRC64;
 Query Match 73.5%; Score 1286; DB 2; Length 478;
 Best Local Similarity 58.9%; Pred. No. 3.4e-75;
 Matches 251; Conservative 38; Mismatches 33; Indels 104; Gaps 2;
 QY 1 MHPQVILSLHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCWNRGSCSLFTCONG 60
 DB 6 MHLQVILSLHLADSVADSVNVDGVAGSIILPCRYNGALISMCHRGICSVFSCPDG 65
 QY 61 IVWTNGTHVYRKDTYKLLGDLSSRDVSLTIENAVSDSGVYCCRVHRGWFNDMKITV 120
 DB 66 IVWTNGTHVYRKDTYKLLGDLSSRDVSLTIENAVSDSGVYCCRVHRGWFNDMKITV 125
 QY 121 SLEIVPPKV----- 129
 DB 126 SLKIGPRVTPIVRTVSTTVPVTTTTLPTTTTLPTTLPTTLPTTLPTTLPTMTT 185
 QY 130 -----TTTTPIVTT 137
 DB 186 LPTTTTVPMTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTL 245
 QY 138 VPTVTIVRTSTVPTTTVPTTTVPTTMSIPTTTVPTTMTVSTTTSVPTTTSIPTTSV 197
 DB 246 LPTTTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTLPTTL 304
 QY 198 PVTTVSTVTPMPLPRQHHEPVATSPSPQAEHTPTTLOGAIRREPTSSPLSYTTDG 257
 DB 305 PTTTWTSTVPTTLPQMDDHEPVATSPSPQAEHTPTTLOGAIRREPTSSPLSYTTDG 364
 QY 258 NDTVTSSDGLMNNNTQLFLEHSLTANTTKGIVAGVCSVLVLLALLGVLIKKYFK 317
 DB 365 SDTVTSSDGLMNNNTQLSPHSPQMNVTTEGIVAGVCSVLVLLALLGVLIKKYFK 424
 QY 318 KEVQQL 323
 DB 425 KEIQQL 430
 RESULT 9
 ID O54947 PRELIMINARY; PRT; 307 AA.
 AC O54947
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)

DE Kidney injury molecule-1 (KIM-1 protein).
 GN Name=KIM-1;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=101116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Sprague-Dawley; TISSUE=Kidney;
 RC MEDLINE=98129827; PubMed=9461608; DOI=10.1074/jbc.273.7.4135;
 RA Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,
 RA Cate R.L., Sanicola M.;
 RT "Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion
 RT molecule containing a novel immunoglobulin domain, is up-regulated in
 RT renal cells after injury.";
 RL J. Biol. Chem. 273:4135-4142(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J.J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Prostate;
 RA Strausberg R.;
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF035963; AAC53546.1; -;
 DR EMBL; BC061820; AAH61820.1; -;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; Ig-like.
 DR SMART; SM00409; IG_1.
 DR PROSITE; PS50835; IG-LIKE; 1.
 SQ SEQUENCE 307 AA; 33963 MW; 736D1DD1F1549760 CRC64;

Query Match 29.9%; Score 522.5; DB 2; Length 307;
 Best Local Similarity 39.4%; Pred. No. 4.2e-26;
 Matches 128; Conservative 39; Mismatches 93; Indels 65; Gaps 12;
 QY 4 QVILSLHLADSVAGSVKVGEGPSVTLPCYSGAVTSMCWNRGSCSLFTCONGI 61
 DB 5 QVIFGLLLLLPGSVDSVEVKGVGHVPTICTYSTYTRGGITTCWGRGQCPYSSQNL 64
 QY 62 VWTNGTHVYRKDTYKLLGDLSSRDVSLTIENAVSDSGVYCCRVHRGWFNDMKITV 121
 DB 65 IWTNGVQVYTRSGRYNIGRISGVDVSLTIENAVSDSGVYCCRVHRGWFNDQWTF 124
 QY 122 LEIVPPKVTTPTVTTVPTTVPVTTVSTTVPVTTTPTTPTTMSIPTTTVPTMTV 179
 DB 125 LE-VKPEIPTSP-----PTRPTTPTTPTTPTTIS----- 153
 QY 180 STTTSVPPTTSTPTTSTVPTTVPVTTTVPVTPMPLPRQHHEPVATSPSPQAEHTPTTLOG 239
 DB 154 TRSTHVPVSTVSTST-----PTPEQTQTHKPEITT-----FYAHEFT--- 191

```

QY 240 AIRREPTSSPLSYTTDGNVTTESSDGLMNNQTLFLEHSLLTANTTKGIYAGVCISV 299
DB 192 ---AEVTETPSYT-PADWNGTWT-SSEAWNNHTVRIPLRP--QRNPTKGFYGVGSVAA 244

QY 300 LVLLALLGVIIAKKY-FPKKVOQL 323
DB 245 LLLLLLASTVTVVTVRIIRKMGSL 269

RESULT 10
Q8VIM2 PRELIMINARY; PRT; 305 AA.
AC Q8VIM2;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DE TIM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c; TISSUE=Spleen;
RX MEDLINE=21582130; PubMed=11725301;
RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Kuchroo V.K.,
RA Barsh G.S., Freeman G.J., Umetsu D.T., DeKruyff R.H.;
RT "Identification of Tapr (an airway hyperreactivity regulatory locus)
and the linked rim gene family.";
RL Nat. Immunol. 2:1109-1116(2001).
DR EMBL; AF399829; AAL35774.1; -.
DR MGD; MGI:2159680; Havcr1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 305 AA; 33391 MW; 8F4EA38627FE85FB CRC64;

Query Match 29.7%; Score 519; DB 2; Length 305;
Best Local Similarity 40.6%; Pred. No. 7e-26;
Matches 129; Conservative 34; Mismatches 95; Indels 60; Gaps 12;

QY 4 QVILSLILHLADSVAGSVKVGEGAPSVTLPCCHYS--GAVTSMCMNRGSCSLFTCQNGI 61
DB 5 QVFISGLLILPGTVDSEVVGKGVGHVPTLPCTYSTVTRGITTTCWGRGQCPSSACQNTL 64
QY 62 VWTNGTHVTRKDRYKLLGDLRRDVSILTENTAVSDGVYCCRVHRGWFNDMKITVS 121
DB 65 IWTNGHRVTVQKSSRYNLKGHISEGDSVLSITIENSVEDSGLYCCRVHPGWFNDQKVTFS 124
QY 122 LEIVPPKVTTPPIVTVPTVTRSTTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
DB 125 LQVKP-----EIPTRPPTTRPTATGRPT-----ISTRTHVPTSRVS 167
QY 161 TTTSVPTTTSIPTTTSVPVTTTSTVTFVPPMPLPRQNHPEVATSPSSQPAPETHPTLQGA 240
DB 168 TSTP-PTSTH--TWTHKEPT---TFCP-----HETAE-- 195
QY 241 IRREPTSSPLSYTTDGNVTTESSDGLMNNQTLFLEHSLLTANTTKGIYAGVCISVL 300
DB 196 VTGIFSHPT----TDWNGTVTSSGD-TWSNHTA--IPPGKPKQKNTKGFYGVGICIAAL 247
QY 301 VLLALLGVIIAKKYFPKK 318
DB 248 LLLLLVSTVAITRYILMK 265

RESULT 11
Q8VIM1 PRELIMINARY; PRT; 282 AA.
AC Q8VIM1;
DT 01-MAR-2002 (TremBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE TIM1.
GN Name=Havcr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2; TISSUE=Spleen;
RX MEDLINE=21582130; PubMed=11725301;
RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Kuchroo V.K.,
RA Barsh G.S., Freeman G.J., Umetsu D.T., DeKruyff R.H.;
RT "Identification of Tapr (an airway hyperreactivity regulatory locus)
and the linked rim gene family.";
RL Nat. Immunol. 2:1109-1116(2001).
DR EMBL; AF399830; AAL35775.1; -.
DR MGD; MGI:2159680; Havcr1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG LIKE; 1.
SQ SEQUENCE 282 AA; 30968 MW; 7D30EE0698F0AC5F CRC64;

Query Match 28.1%; Score 491.5; DB 2; Length 282;
Best Local Similarity 36.6%; Pred. No. 3.9e-24;
Matches 116; Conservative 32; Mismatches 88; Indels 81; Gaps 7;

QY 4 QVILSLILHLADSVAGSVKVGEGAPSVTLPCCHYS--GAVTSMCMNRGSCSLFTCQNGI 61
DB 5 QVFISGLLILPGAVDSVVGKGVGHVPTLPCTYSTVTRGITTTCWGRGQCPSSACQNTL 64
QY 62 VWTNGTHVTRKDRYKLLGDLRRDVSILTENTAVSDGVYCCRVHRGWFNDMKITVS 121
DB 65 IWTNGHRVTVQKSSRYNLKGHISEGDSVLSITIENSVEDSGLYCCRVHPGWFNDQKVTFS 124
QY 122 LEIVPPKVTTPPIVTVPTVTRSTTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
DB 125 LQVKP-----EIPTRPPTTRPTATGRPT-----TIST 155
QY 182 TTTSVPTTTSIPTTTSVPVTTTSTVTFVPPMPLPRQNHPEVATSPSSQPAPETHPTLQGA 241
DB 156 -----RSTHVPSTSTVSTP----- 171
QY 242 IRREPTSSPLSYTTDGNVTTESSDGLMNNQTLFLEHSLLTANTTKGIYAGVCISVLV 301
DB 172 ---PTSTHTWTHKPDWNGTVTSSGD-TWSNHTA--IPPGKPKQKNTKGFYGVGICIAALL 225
QY 302 LLLALLGVIIAKKYFPKK 318
DB 226 LLLLVSTVAITRYILMK 242

RESULT 12
Q7TPU2 PRELIMINARY; PRT; 282 AA.
AC Q7TPU2;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Blastocyst;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

```


RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Murny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J.C., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smalish J.E., Schmerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2].
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Blastocyst;
 RA Strausberg R.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC053400; AAH53400.1; -;
 DR InterPro; IPR003599; IG-like.
 DR SMART; IPR007110; IG-like.
 DR PROSITE; PS00409; IG; 1.
 DR PROSITE; PS0835; IG LIKE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 282 AA; 30938 MW; 6D31E2778480AC5F CRC64;
 Query Match 27.8%; Score 486.5; DB 2; Length 282;
 Best Local Similarity 36.3%; Pred. No. 8.1e-24;
 Matches 115; Conservative 32; Mismatches 89; Indels 81; Gaps 7;
 QY 4 QVILSLILHLADSVAGSVKVGEGAPSVTLPCYHS--GAVTSMCNWRCGSLFTQNGI 61
 DB 5 QVFISGLILLLPGAVDSYEVGVGHPVTLPTCTYRGTITTCWGRQCPSACQNTL 64
 QY 62 VWTNGTHVYRKDTRYKLLGLDLSRRDVSLLTIENTAVSDSGVYCCRVHRGWFNDKMTVS 121
 DB 65 IWTNGHRTYQKSRYNLKHISGDSVLSIENSVESDGLCCRVIEFGWENDQKVTFS 124
 QY 122 LEIVPPKVTTPITVPTVTRSTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 181
 DB 125 LQVKP-----EIPTRPPRRPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 155
 QY 182 TTSVPTTTSIPTTSVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 241
 DB 156 -----RSTHPTSTRVSTSTP-----PSSPOPAETHPTT 171
 QY 242 RREPTSSPLSYTTDGNQDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVLV 301
 DB 172 ----PSTHTWTHKPDWNGVTSSGD--TWSNHTA--IPPKQKQKPTRGFYVGCIAL 225
 QY 302 LLALLGVIIAKYFFKK 318
 DB 226 LLLLVSTVAITRYILMK 242
 RESULT 13
 Q8VBWO PRELIMINARY; PRT; 305 AA.
 AC Q8VBWO;
 DT 01-MAR-2002 (Tremblrel. 20, Created)
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE TIM2.
 GN Name=Tim2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2, and BALB/c; TISSUE=Spleen;
 RX MEDLINE=21582130; PubMed=11725301;
 RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Kuchroo V.K.,
 RA Barsh G.S., Freeman G.J., Umetsu D.T., Dekruyff R.H.;
 RT "Identification of Tapr (an airway hyperreactivity regulatory locus)
 RT and the linked Tm gene family.";
 RL Nat. Immunol. 2:1109-1116 (2001).
 DR EMBL; AF399828; AAL35773.1; -;
 DR EMBL; AF399827; AAL35772.1; -;
 DR MGD; MGI:2159681; Timd2.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR SMART; SM00409; IG; 1.
 DR PROSITE; PS0835; IG LIKE; 1.
 SQ SEQUENCE 305 AA; 33519 MW; 68B69370FE23C9D0 CRC64;
 Query Match 22.9%; Score 401; DB 2; Length 305;
 Best Local Similarity 33.5%; Pred. No. 2.9e-18;
 Matches 108; Conservative 48; Mismatches 92; Indels 74; Gaps 12;
 QY 4 QVILSLILHLADSVAGSVKVGEGAPSVTLPCYHS--GAVTSMCNWRCGSLFTQNG 60
 DB 5 QVFISGLILLLPGAVESHTAVOGLAGHPVTLPCIVSTHLGGIVPMCWGLGECRHSYCI 64
 QY 61 IVWTNGTHVYRKDTRYKLLGLDLSRRDVSLLTIENTAVSDSGVYCCRVHRGWFNDKMT 120
 DB 65 LIWTNGYTVTHQRNSRYQLKGNISGNSVLSIENVTGDDGPGVCCVVEIPGAPH--FVD 122
 QY 121 SLIEIVPPKVTTPITVPTVTRSTVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 180
 DB 123 MLE-VKPEISTSP-----PTR-----PTATGRPTTIS-T 149
 QY 181 TTSVPTTTSIPTTSVPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT 236
 DB 150 RSTHPTSTRVSTSTP-----PTAHTETIKPEATFTPDQTAETELPST 197
 QY 237 LQGAIRREPTSSPLSYTTDGNQDVTSSDGLWNNQTLFLEHSLLTANTTKGIYAGVC 296
 DB 198 -----PADWNTVT--SSDDPDWDN--TEV--IPPKQKQKLNKGFVVGIS 237
 QY 297 ISVLVLLALLGVIIAKYFFKK 318
 DB 238 IAALLILLLSTMTVITRYVVK 259
 RESULT 14
 Q8CIC7 PRELIMINARY; PRT; 362 AA.
 AC Q8CIC7;
 DT 01-MAR-2003 (Tremblrel. 23, Created)
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
 DE T-cell immunoglobulin and mucin domain containing 4.
 GN Name=Timd4;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler N.K.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 3.79568 Seconds
(without alignments)
1064.659 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225

Sequence: 1 PLYSYTTCGNDVTETSSDGL.....NNQTQLFLEHSLLTANTTKG 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: Piri:*

2: Piri:*

3: Piri:*

4: Piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	178	79.1	451	2 S71754	cellular hepatitis
2	62.5	27.8	639	2 S48959	mitochondrial oute
3	61.5	27.3	1269	2 A90267	proteinase related
4	60	26.7	602	2 A45769	acetylcholine rece
5	59	26.2	1254	1 JQ1978	structural polypro
6	58	25.8	1254	1 VHWVVE	structural polypro
7	58	25.8	1254	1 VHWVVT	structural polypro
8	58	25.8	1255	1 B44213	structural polypro
9	58	25.8	1255	1 D44213	structural polypro
10	57.5	25.6	405	2 H89930	protein Rllgl.14
11	57	25.3	2399	2 T71879	toxin-like outer m
12	56.5	25.1	326	1 VGXRER	glycoprotein Vp7 p
13	56.5	25.1	338	2 S75217	N-acetylmuramyl-L
14	56.5	25.1	1658	2 D86890	DNA-directed DNA p
15	56	24.9	472	2 S36519	I2 protein - human
16	55.5	24.7	322	2 T23891	hypothetical prote
17	55	24.4	826	1 T02753	S-receptor kinase
18	55	24.4	846	2 C82135	chitinase VC1952 I
19	54.5	24.2	189	2 B26659	T-cell receptor ga
20	54.5	24.2	205	2 B26425	T-cell receptor ga
21	54.5	24.2	213	2 T07115	hypothetical prote
22	54.5	24.2	285	2 G85016	probable myb-relat
23	54.5	24.2	304	2 S04663	T-cell receptor ga
24	54.5	24.2	323	2 S01895	T-cell receptor ga
25	54.5	24.2	386	2 A82284	conserved hypothet
26	54.5	24.2	639	2 A55019	muscarinic acetyl
27	54.5	24.2	923	2 E70820	hypothetical glyci
28	54	24.0	735	2 B83763	catalase BH0906 li
29	54	24.0	891	2 T30812	ubiquitin-protein

ALIGNMENTS

RESULT 1

S71754

cellular hepatitis A receptor HAVcr-1 precursor - green monkey

N:Alternate names: surface glycoprotein

C:Species: Cercopithecus aethiops (green monkey, grivet)

C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004

C:Accession: S71754

R:Kaplan, G.; Totsuka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstein, S.M.

EMBO J. 15, 4282-4296, 1996

A:Title: Identification of a surface glycoprotein on African green monkey kidney cells as

A:Reference number: S71754; MUID:97015129; PMID:8861957

A:Accession: S71754

A:Molecule type: mRNA

A:Residues: 1-451 <KAP>

A:Cross-references: UNIPROT:Q95144; EMBL:X98252; NID:gl526573; PID:e247449; PID:gl526574

A:Experimental source: kidney

C:Keywords: glycoprotein

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match 79.1%; Score 178; DB 2; Length 451;

Best Local Similarity 81.0%; Pred. No. 1.7e-14;

Matches 34; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTTCGNDVTETSSDGLMNNQTLFLEHSLLTANTTKG 42

DB 329 PLYSYTTCGNDVTETSSDGLMNNQTLSPHSQWVNTTEG 370

RESULT 2

S48959

mitochondrial outer membrane protein TOM71 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YHR117W

C:Species: Saccharomyces cerevisiae

C>Date: 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change 09-Jul-2004

C:Accession: S48959; S62886

R:Latreille, P.

submitted to the EMBL Data Library, May 1994

A:Description: The sequence of S. cerevisiae cosmid 8263.

A:Reference number: S46676

A:Accession: S48959

A:Molecule type: DNA

A:Residues: 1-639 <LAT>

A:Cross-references: UNIPROT:P38825; EMBL:U00059; NID:gs529116; PIDN:AAB68868.1; PID:gs52911

R:Boemer, U.; Pfanner, N.; Dietmeier, K.

FEBS Lett. 382, 153-158, 1996

A:Title: Identification of a third yeast mitochondrial Tom protein with tetratricopeptid

A:Reference number: S62886; MUID:96196584; PMID:8612740

A:Accession: S62886

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-639 <BOE>

C:Genetics:

A:Gene: SGD.TOM71; TOM72
A:Cross-references: SGD:S0001159; MIPS:YHR117W
A:Map position: 8R

A:Genome: nuclear

C:Superfamily: mitochondrial outer membrane protein, 70K; tetratricopeptide repeat homolog
F:86-119/Domain: tetratricopeptide repeat homolog <TT01>
F:127-160/Domain: tetratricopeptide repeat homolog <TT02>
F:161-194/Domain: tetratricopeptide repeat homolog <TT03>
F:345-377/Domain: tetratricopeptide repeat homolog <TT04>
F:378-411/Domain: tetratricopeptide repeat homolog <TT05>
F:412-445/Domain: tetratricopeptide repeat homolog <TT06>
F:446-479/Domain: tetratricopeptide repeat homolog <TT07>
F:480-513/Domain: tetratricopeptide repeat homolog <TT08>
F:530-563/Domain: tetratricopeptide repeat homolog <TT09>
F:564-597/Domain: tetratricopeptide repeat homolog <TT10>

Query Match 27.8%; Score 62.5; DB 2; Length 639;

Best Local Similarity 40.5%; Pred. No. 6.8;
Matches 17; Conservative 6; Mismatches 4; Indels 15; Gaps 2;

QY 2 LYSYTTDG----NDVTSSDGLWNNNTQFLFHSLLTANT 39

DB 305 LYSATDEGYLVANLLTKSTD-----MYHSLLSANT 335

RESULT 3

A:Proteinase related protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: A90267

R:Singh, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thai-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: A90267

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1269 <KUR>

A:Cross-references: UNIPROT:Q97206; GB:AE006641; NID:g13814328; PIDN:AAK41392.1; GSPDB:G

C:Genetics:

A:Gene: SSO1141

Query Match 27.3%; Score 61.5; DB 2; Length 1269;

Best Local Similarity 34.0%; Pred. No. 20;

Matches 18; Conservative 7; Mismatches 11; Indels 17; Gaps 2;

QY 4 SYTTDG-----NDVTSSDGLWNNNTQFLFHSLLTANT 39

DB 1163 SPTTNGTHIVTNTQYPDGGYNLTATQSDGLSSNSSVLYFENGTLNLT 1215

RESULT 4

A:45769

A:Title: acetylcholine receptor synthesis stimulator ARIA-1 precursor - chicken

C:Species: Gallus gallus (chicken)

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Jul-2004

C:Accession: A45769

R:Fallis, D.L.; Rosen, K.M.; Corfas, G.; Lane, W.S.; Fischbach, G.D.

Cell 72, 801-815, 1993

A:Title: ARIA, a protein that stimulates acetylcholine receptor synthesis, is a member of

A:Reference number: A45769; MUID:93201602; PMID:8453670

A:Accession: A45769

A:Status: preliminary

A:Molecule type: mRNA; protein

A:Residues: 1-602 <FAL>

A:Cross-references: UNIPROT:Q05199; GB:I11264; NID:g212603; PIDN:AAA49037.1; PID:g212604

A:Experimental source: brain

A>Note: sequence extracted from NCBI backbone (NCBIN:127787, NCBI:P:127788)

C:Superfamily: human heregulin; EGF homology; immunoglobulin homology

Query Match 26.7%; Score 60; DB 2; Length 602;

Best Local Similarity 31.7%; Pred. No. 13;

Matches 13; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 3 YSYTTDGDNTVTSSDGLWNNNTQFLF--EHSLLTANTTK 41

DB 301 YTSYTHSWTWTQTSPSHWSNGHTESILSESHSVLVSSSV 341

RESULT 5

JQ1978

structural polyprotein - Venezuelan equine encephalitis virus (subtype II, strain Everglades

N:Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2;

C:Species: Venezuelan equine encephalitis virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004

C:Accession: JQ1978

R:Snider, J.M.; Kinney, R.M.; Tsuchiya, K.R.; Trent, D.W.

J. Gen. Virol. 74, 519-523, 1993

A:Title: Molecular evidence that epizootic Venezuelan equine encephalitis (VEE) I-AB viru

A:Reference number: JQ1978; MUID:93187617; PMID:8445371

A:Accession: JQ1978

A:Molecule type: mRNA

A:Residues: 1-1254 <SNE>

A:Cross-references: UNIPROT:P36330; GB:L04598; NID:g290612; PIDN:AAA42984.1; PID:g290614

C:Superfamily: togavirus structural polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-274/Product: coat protein #status predicted <CTP>

F:275-333/Product: membrane glycoprotein E3 #status predicted <MG3>

F:334-756/Product: membrane glycoprotein E2 #status predicted <MG2>

F:701-721/Domain: transmembrane #status predicted <TM1>

F:751-812/Product: 6K protein #status predicted <KP6>

F:794-813/Domain: transmembrane #status predicted <TM2>

F:813-1254/Product: membrane glycoprotein E1 #status predicted <MG1>

F:1231-1248/Domain: transmembrane #status predicted <TM3>

F:47,285,545,651,946/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 26.2%; Score 59; DB 1; Length 1254;

Best Local Similarity 45.5%; Pred. No. 40;

Matches 15; Conservative 3; Mismatches 11; Indels 4; Gaps 1;

QY 11 DTVTSSDGLWNNNO----TOLFLEHSLLTANT 39

DB 757 ETTWESLDHLWNNNQMFWTQLPLAALIVVT 789

RESULT 6

VHWVVR

structural polyprotein - Venezuelan equine encephalitis virus (strain TC-83)

N:Contains: 6K peptide; coat protein; membrane glycoprotein E1; membrane glycoprotein E2;

C:Species: Venezuelan equine encephalitis virus

C:Date: 30-Sep-1988 #sequence_revision 30-Sep-1988 #text_change 09-Jul-2004

C:Accession: A27871

R:Johnson, B.J.B.; Kinney, R.M.; Kost, C.L.; Trent, D.W.

J. Gen. Virol. 67, 1951-1960, 1986

A:Title: Molecular determinants of alphavirus neurovirulence: nucleotide and deduced prot

A:Reference number: A27871; MUID:86306669; PMID:3755750

A:Accession: A27871

A:Molecule type: mRNA

A:Residues: 1-1254 <JOH>

A:Cross-references: UNIPROT:P05674; GB:X04368; NID:g58621; PIDN:CAA27883.1; PID:g58622

A>Note: the authors translated the codon UUC for residues 19 and 27 as Ile and ACC for r

C:Superfamily: togavirus structural polyprotein

C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein

F:1-275/Product: coat protein #status predicted <COP>

F:276-334/Product: membrane glycoprotein E3 #status predicted <GP3>

F:335-757/Product: membrane glycoprotein E2 #status predicted <GP2>

F:702-718/Domain: transmembrane #status predicted <TM1>

F:758-812/Product: 6K protein #status predicted <P1P>

F:774-790/Domain: transmembrane #status predicted <TM2>

F:792-808/Domain: transmembrane #status predicted <TM3>

F:813-1254/Product: membrane glycoprotein E1 #status predicted <GP1>

```

F:1231-1247/Domain: transmembrane #status predicted <TM4>
F:47,286,546,625,652,946/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 25.8%; Score 58; DB 1; Length 1254;
Best Local Similarity 47.6%; Pred. No. 54;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 11 DVTWSSDGLWNNNTQLFLE 31
DB 758 ETTWESLDHLWNNNQMFQ 778
:::|||||:::
:::|||||:::

RESULT 7
VHWVVT
structural polyprotein - Venezuelan equine encephalitis virus (strain TRD)
N:Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: Venezuelan equine encephalitis virus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: B31467; A47612
R:Kinney, R.M.; Johnson, B.J.B.; Welch, J.B.; Teuchiya, K.R.; Trent, D.W.
Virology 170, 19-30, 1989
A:Title: The full-length nucleotide sequences of the virulent Trinidad donkey strain of
A:Reference number: A31467; MUID:69243175; PMID:2524126
A:Accession: B31467
A:Molecule type: mRNA
A:Residues: 1-1254 <K1>
A:Cross-references: UNIPROT:P09592; GB:J04332; NID:g232708; PIDN:AAB02519.1; PID:g323710
R:Kinney, R.M.; Johnson, B.J.B.; Brown, V.L.; Trent, D.W.
Virology 152, 400-413, 1986
A:Title: Nucleotide sequence of the 26 S mRNA of the virulent Trinidad donkey strain of
A:Reference number: A47612; MUID:86263392; PMID:3088830
A:Accession: A47612
A:Molecule type: mRNA
A:Residues: 1-542, 'K', 544-810, 'P', 812-1254 <K12>
A:Cross-references: GB:L01442
C:Superfamily: togavirus structural polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-275/Product: coat protein #status predicted <CTP>
F:276-334/Product: membrane glycoprotein E3 #status predicted <MG3>
F:335-757/Product: membrane glycoprotein E2 #status predicted <MG2>
F:701-718/Domain: transmembrane #status predicted <TM1>
F:758-812/Product: 6K protein #status predicted <KP6>
F:774-790/Domain: transmembrane #status predicted <TM2>
F:795-813/Domain: transmembrane #status predicted <TM3>
F:813-1254/Product: membrane glycoprotein E1 #status predicted <MG1>
F:1231-1248/Domain: transmembrane #status predicted <TM4>
F:47,286,546,652,946/Binding site: carbohydrate (Asn) (covalent) #status predicted
Query Match 25.8%; Score 58; DB 1; Length 1254;
Best Local Similarity 47.6%; Pred. No. 54;
Matches 10; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 11 DVTWSSDGLWNNNTQLFLE 31
DB 758 ETTWESLDHLWNNNQMFQ 778
:::|||||:::
:::|||||:::

RESULT 8
B44213
structural polyprotein - Venezuelan equine encephalitis virus (strain P676)
N:Contains: 6K protein; coat protein; membrane glycoprotein E1; membrane glycoprotein E2
C:Species: Venezuelan equine encephalitis virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C:Accession: B44213
R:Kinney, R.M.; Teuchiya, K.R.; Snieder, J.M.; Trent, D.W.
Virology 191, 569-580, 1992
A:Title: Genetic evidence that epizootic Venezuelan equine encephalitis (VEE) viruses are
A:Reference number: A44213; MUID:93079859; PMID:1448915
A:Accession: B44213
A:Molecule type: genomic RNA
A:Residues: 1-1255 <K1N>
A:Cross-references: UNIPROT:P36332; GB:L04653; NID:g290609; PIDN:AAC19319.1; PID:g290611
C:Superfamily: togavirus structural polyprotein

```



```

R;Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A;Description: Primer-directed sequencing of human papillomavirus types.
A;Reference number: S36469
A;Accession: S36519
A;Molecule type: DNA
A;Residues: 1-472 <DEL>
A;Cross-references: UNIPROT:P36758; EMBL:X74476; NID:G396989; PIDN:CAA52559.1; PID:G3969
C;Superfamily: papillomavirus L2 protein
C;Keywords: late protein

Query Match          24.9%; Score 56; DB 2; Length 472;
Best Local Similarity 38.5%; Pred. No. 31;
Matches 15; Conservative 2; Mismatches 22; Indels 0; Gaps 0;

Qy      2  LYSYTTDGDNDVTSSDGLWNNNQTLQFLFHSLLITANTT 40
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      363  LYDVLDDNVDITEVETPTGTNTQSVFASEISITANTT 401

Search completed: June 29, 2005, 08:58:12
Job time : 4.79568 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:46:27 ; Search time 49.0138 Seconds
(without alignments)
635.503 Million cell updates/sec

Title: US-10-718-321-6

Perfect score: 432

Sequence: 1 MPFPRQNHBPVATSPSPQP.....NNOTQFLFLESLTANTTKG 81

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 487488

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	22.9	18	17	US-10-718-321-1
2	72	16.7	14	17	US-10-805-177-87
3	67	15.5	12	17	US-10-805-177-92
4	63	14.6	12	17	US-10-805-177-88
5	60	13.9	50	9	US-09-864-761-40317
6	57	13.2	45	15	US-10-424-599-258666
7	56	13.0	10	17	US-10-805-177-93
8	56	13.0	47	16	US-10-425-115-195346
9	54.5	12.6	40	16	US-10-481-180-489
10	54.5	12.6	47	16	US-10-425-115-358602
11	52.5	12.2	47	10	US-09-986-480-394
					Sequence 1, Appli
					Sequence 87, Appl
					Sequence 92, Appl
					Sequence 88, Appl
					Sequence 40317, A
					Sequence 258666,
					Sequence 93, Appl
					Sequence 195346,
					Sequence 489, App
					Sequence 358602,
					Sequence 394, App

12	52.5	12.2	47	16	US-10-425-115-354965
13	52.5	12.2	47	17	US-10-863-332-394
14	52.5	12.2	48	15	US-10-424-599-162113
15	52	12.0	10	17	US-10-805-177-89
16	52	12.0	47	14	US-10-029-386-32839
17	51.5	11.9	50	10	US-09-933-767-540
18	51.5	11.9	50	14	US-10-004-860-540
19	51.5	11.9	50	14	US-10-023-282-540
20	50.5	11.7	47	15	US-10-424-599-272582
21	50.5	11.7	47	15	US-10-424-599-272582
22	48.5	11.2	41	16	US-10-425-115-312293
23	48.5	11.2	42	15	US-10-424-599-283392
24	48.5	11.2	45	16	US-10-425-115-232573
25	48.5	11.2	48	14	US-10-008-524A-164
26	48.5	11.2	48	15	US-10-350-719-164
27	48	11.1	44	15	US-10-424-599-190016
28	48	11.1	45	16	US-10-425-115-223316
29	48	11.1	48	14	US-10-029-386-29088
30	48	11.1	48	15	US-10-424-599-28496
31	48	11.1	49	15	US-10-424-599-192216
32	48	11.1	50	16	US-10-425-115-256042
33	47.5	11.0	40	15	US-10-424-599-264613
34	47.5	11.0	40	16	US-10-425-115-352983
35	47.5	11.0	46	16	US-10-425-115-203986
36	47.5	11.0	50	11	US-09-864-408A-8636
37	47	10.9	8	17	US-10-805-177-104
38	47	10.9	8	17	US-10-805-177-106
39	47	10.9	10	17	US-10-805-177-98
40	47	10.9	23	16	US-10-481-180-474
41	47	10.9	47	16	US-10-437-963-174713
42	47	10.9	47	16	US-10-425-115-188216
43	46.5	10.8	34	16	US-10-425-115-309663
44	46.5	10.8	44	15	US-10-424-599-208981
45	46.5	10.8	46	10	US-09-764-891-5368

ALIGNMENTS

RESULT 1
US-10-718-321-1
; Sequence 1, Application US/10718321
; Publication No. US2005012117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE REFERENCE: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-1

Query Match 22.9%; Score 99; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 55 SSGLWNNNTQFLFLEHS 72
|||||

Db 1 SSDGLMNNQTQLFLEHS 18

RESULT 2

US-10-805-177-87
; Sequence 87, Application US/10805177
; Publication No. US20050084449A1
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Chen, Francine
; APPLICANT: Bezabeh, Binyam
; APPLICANT: Foltz, Ian
; APPLICANT: Tse, Kam Fai
; APPLICANT: Jeffers, Michael
; APPLICANT: Mesri, Mehdi
; APPLICANT: Starling, Gary
; APPLICANT: Mezes, Peter
; APPLICANT: Khrantsov, Nikolai
; TITLE OF INVENTION: ANTIBODIES AGAINST T CELL IMMUNOGLOBULIN
; TITLE OF INVENTION: DOMAIN AND MUCIN DOMAIN 1 (TIM-1) ANTIGEN AND USES THEREOF
; FILE REFERENCE: ABXCUR.006A
; CURRENT APPLICATION NUMBER: US/10/805,177
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/456,652
; PRIOR FILING DATE: 2003-03-19
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-805-177-87

Query Match 16.7%; Score 72; DB 17; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHEPVAT 13

Db 2 MPLPRQNHEPVAT 14

RESULT 3

US-10-805-177-92
; Sequence 92, Application US/10805177
; Publication No. US20050084449A1
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Chen, Francine
; APPLICANT: Bezabeh, Binyam
; APPLICANT: Foltz, Ian
; APPLICANT: Tse, Kam Fai
; APPLICANT: Jeffers, Michael
; APPLICANT: Mesri, Mehdi
; APPLICANT: Starling, Gary
; APPLICANT: Mezes, Peter
; APPLICANT: Khrantsov, Nikolai
; TITLE OF INVENTION: ANTIBODIES AGAINST T CELL IMMUNOGLOBULIN
; TITLE OF INVENTION: DOMAIN AND MUCIN DOMAIN 1 (TIM-1) ANTIGEN AND USES THEREOF
; FILE REFERENCE: ABXCUR.006A
; CURRENT APPLICATION NUMBER: US/10/805,177
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/456,652
; PRIOR FILING DATE: 2003-03-19
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-805-177-92

Query Match 15.5%; Score 67; DB 17; Length 12;

Best Local Similarity 100.0%; Pred. No. 3;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PLPRQNHEPVAT 13

Db 1 PLPRQNHEPVAT 12

RESULT 4

US-10-805-177-88
; Sequence 88, Application US/10805177
; Publication No. US20050084449A1
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Chen, Francine
; APPLICANT: Bezabeh, Binyam
; APPLICANT: Foltz, Ian
; APPLICANT: Tse, Kam Fai
; APPLICANT: Jeffers, Michael
; APPLICANT: Mesri, Mehdi
; APPLICANT: Starling, Gary
; APPLICANT: Mezes, Peter
; APPLICANT: Khrantsov, Nikolai
; TITLE OF INVENTION: ANTIBODIES AGAINST T CELL IMMUNOGLOBULIN
; TITLE OF INVENTION: DOMAIN AND MUCIN DOMAIN 1 (TIM-1) ANTIGEN AND USES THEREOF
; FILE REFERENCE: ABXCUR.006A
; CURRENT APPLICATION NUMBER: US/10/805,177
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/456,652
; PRIOR FILING DATE: 2003-03-19
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-805-177-88

Query Match 14.6%; Score 63; DB 17; Length 12;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHEPV 11

Db 2 MPLPRQNHEPV 12

RESULT 5

US-09-864-761-40317
; Sequence 40317, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

RESULT 6
US-10-424-599-258666
; Sequence 258666, Application US/10424599
; Publication No. US20040031072A1
;
GENERAL INFORMATION:
;
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
;
; TITLE OF INVENTION: Soy Nucleic Acid Mo
; TITLE OF INVENTION: Plants and Uses Th
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,5
; CURRENT FILING DATE: 2003-04-28
;
; NUMBER OF SEQ ID NOS: 285684
;
; SEQ ID NO 258666
;
; LENGTH: 45
;
; TYPE: PRT
;
; ORGANISM: Glycine max
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT MRT38

```

US-10-425-115-195346

US-10-425-115-195346

OTHER INFORMATION: CLOHE ID: MK143771
 NS-10-425-115-105246

Matches 20; Conservative 4; Mismatches 16; Indels 7; Gaps 3;

Qy 13 TSPSPQAPAE--THPTTLOGAIRREPTSSPL-----YSTTDGNDTDT 53
| | | | | : | | | | | : | | | | | : | | | | | :
Db 1 TSMSTTGPSETTTSPTTALTMSR-PTSMPLTWQKSYITISKFTVS 46
| | | | | : | | | | | : | | | | | : | | | | | :

RESULT 13

US-10-863-332-394
; Sequence 394, Application US/10863332
; Publication No. US20050064458A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 143 Human Secreted Proteins
; FILE REFERENCE: P5500P1
; CURRENT APPLICATION NUMBER: US/10/863,332
; CURRENT FILING DATE: 2004-06-09
; PRIOR APPLICATION NUMBER: US/09/986,480
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: PCT/US00/12788
; PRIOR FILING DATE: 2000-05-11
; PRIOR APPLICATION NUMBER: US 60/134,068
; PRIOR FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 456
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 394
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (37)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-863-332-394

Query Match 12.2%; Score 52.5; DB 17; Length 47;
Best Local Similarity 35.3%; Pred. No. 4.6e+02;
Matches 12; Conservative 2; Mismatches 13; Indels 7; Gaps 1;

Qy 7 NHEPVATSPSPQAPAEHTPTTLOGAIRREPTSSP 40
| | | | | : | | | | | : | | | | | : | | | | | :
Db 2 NMNPVSTLPLGSPSPSHTA-----HPPTPSP 28
| | | | | : | | | | | : | | | | | : | | | | | :

RESULT 14

US-10-424-599-162113
; Sequence 162113, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 162113
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_117405C.1.pep
US-10-424-599-162113

Query Match 12.2%; Score 52.5; DB 15; Length 48;
Best Local Similarity 31.5%; Pred. No. 4.8e+02;

Matches 17; Conservative 3; Mismatches 27; Indels 7; Gaps 1;

Qy 5 RQNEHPVATSPSPQAPAEHTPTTLOGAIRREPTSSPLSYTTDGNDTVTSSDG 58
| | | | | : | | | | | : | | | | | : | | | | | :
Db 1 RDXHEPVSPSPNSAAMS-----LGTINLVTFSSAQSRMVSGVDVTAHPSTG 47
| | | | | : | | | | | : | | | | | : | | | | | :

RESULT 15

US-10-805-177-89
; Sequence 89, Application US/10805177
; Publication No. US20050084449A1
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Chen, Francine
; APPLICANT: Bezabeh, Binyam
; APPLICANT: Foltz, Ian
; APPLICANT: Tse, Kam Fai
; APPLICANT: Jeffers, Michael
; APPLICANT: Mesiri, Mehdi
; APPLICANT: Starling, Gary
; APPLICANT: Mezes, Peter
; APPLICANT: Khrantsov, Nikolai
; TITLE OF INVENTION: ANTIBODIES AGAINST T CELL IMMUNOGLOBULIN
; TITLE OF INVENTION: DOMAIN AND MUCIN DOMAIN 1 (TIM-1) ANTIGEN AND USES THEREOF
; FILE REFERENCE: ABXCUR.006A
; CURRENT APPLICATION NUMBER: US/10/805,177
; CURRENT FILING DATE: 2004-03-19
; PRIOR APPLICATION NUMBER: 60/456,652
; PRIOR FILING DATE: 2003-03-19
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 89
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-805-177-89

Query Match 12.0%; Score 52; DB 17; Length 10;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLPRQNH 9
| | | | | : | | | | | : | | | | | : | | | | | :
Db 2 MPLPRQNH 10
| | | | | : | | | | | : | | | | | : | | | | | :

Search completed: June 29, 2005, 10:14:18
Job time : 49.0138 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 11.1395 Seconds
(without alignments)
542.805 Million cell updates/sec

Title: US-10-718-321-6
Perfect score: 432
Sequence: 1 MBLPRQNHVPATSPSPQP.....NNQTQLFLEHSLLTANTTKG 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432	100.0	334	US-09-197-970B-7	Sequence 7, Appli
2	348	80.6	451	US-08-287-001A-2	Sequence 2, Appli
3	348	80.6	451	PCT-US95-09941-2	Sequence 2, Appli
4	89	20.6	307	US-09-197-970B-3	Sequence 3, Appli
5	71.5	16.6	297	US-09-006-535-4	Sequence 4, Appli
6	71	16.4	1911	US-09-854-856-64	Sequence 64, Appli
7	71	16.4	1939	US-09-854-856-48	Sequence 48, Appli
8	71	16.4	1971	US-09-854-856-32	Sequence 32, Appli
9	71	16.4	1999	US-09-854-856-16	Sequence 16, Appli
10	71	16.4	2004	US-09-854-856-58	Sequence 58, Appli
11	71	16.4	2032	US-09-854-856-42	Sequence 42, Appli
12	71	16.4	2048	US-09-854-856-62	Sequence 62, Appli
13	71	16.4	2064	US-09-854-856-26	Sequence 26, Appli
14	71	16.4	2076	US-09-854-856-46	Sequence 46, Appli
15	71	16.4	2092	US-09-854-856-10	Sequence 10, Appli
16	71	16.4	2108	US-09-854-856-30	Sequence 30, Appli
17	71	16.4	2136	US-09-854-856-14	Sequence 14, Appli
18	71	16.4	2141	US-09-854-856-56	Sequence 56, Appli
19	71	16.4	2157	US-09-854-856-52	Sequence 52, Appli
20	71	16.4	2169	US-09-854-856-40	Sequence 40, Appli
21	71	16.4	2185	US-09-854-856-36	Sequence 36, Appli
22	71	16.4	2201	US-09-854-856-24	Sequence 24, Appli
23	71	16.4	2217	US-09-854-856-20	Sequence 20, Appli
24	71	16.4	2229	US-09-854-856-8	Sequence 8, Appli
25	71	16.4	2245	US-09-854-856-4	Sequence 4, Appli
26	71	16.4	2294	US-09-854-856-50	Sequence 50, Appli
27	71	16.4	2322	US-09-854-856-34	Sequence 34, Appli

28	71	16.4	2354	4	US-09-854-856-18	Sequence 18, Appli
29	71	16.4	2382	4	US-09-854-856-2	Sequence 2, Appli
30	70	16.2	109	4	US-09-270-767-46395	Sequence 46395, A
31	69.5	16.1	196	4	US-09-248-796A-17562	Sequence 17562, A
32	69.5	16.1	445	2	US-08-900-148-2	Sequence 2, Appli
33	69.5	16.1	479	4	US-09-248-796A-20593	Sequence 20593, A
34	69	16.0	567	4	US-09-514-245-6	Sequence 6, Appli
35	69	16.0	602	1	US-08-168-091A-2	Sequence 2, Appli
36	69	16.0	602	1	US-08-428-926-5	Sequence 5, Appli
37	69	16.0	602	1	US-08-428-927-5	Sequence 5, Appli
38	69	16.0	602	1	US-08-428-298-5	Sequence 5, Appli
39	69	16.0	602	1	US-08-339-517-5	Sequence 5, Appli
40	69	16.0	1070	3	US-08-697-954-2	Sequence 2, Appli
41	69	16.0	1166	4	US-09-200-650E-7	Sequence 7, Appli
42	68.5	15.9	319	4	US-09-270-767-44343	Sequence 44343, A
43	68.5	15.9	681	3	US-08-760-615-4	Sequence 4, Appli
44	68.5	15.9	681	4	US-09-336-910A-2	Sequence 2, Appli
45	67.5	15.6	160	4	US-09-270-767-46213	Sequence 46213, A

ALIGNMENTS

RESULT 1
US-09-197-970B-7
; Sequence 7, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-No. 6664385-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-197-970B-7

Query Match 100.0%; Score 432; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. NO. 2e-41;

```
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MFLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 60
Db 210 MFLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 269
QY 61 NNNQTLFLEHSLLTANTTKG 81
Db 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 2
US-08-287-001A-2
; Sequence 2, Application US/08287001A
; Patent No. 5622861
; GENERAL INFORMATION:
; APPLICANT: KAPLAN, GERARDO
; APPLICANT: FEINSTONE, STEPHEN M.
; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,001A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09941-2

Query Match 80.6%; Score 348; DB 1; Length 451;
Best Local Similarity 82.5%; Pred. No. 1.2e-31;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 PLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 61
Db 291 PLPMQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 350
QY 62 NNNQTLFLEHSLLTANTTKG 81
Db 351 NNNQTLSPHSPQMVNTTEG 370

US-08-287-001A-2

Query Match 80.6%; Score 348; DB 1; Length 451;
Best Local Similarity 82.5%; Pred. No. 1.2e-31;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 PLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 61
Db 291 PLPMQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 350
QY 62 NNNQTLFLEHSLLTANTTKG 81
Db 351 NNNQTLSPHSPQMVNTTEG 370

US-08-287-001A-2

Query Match 80.6%; Score 348; DB 5; Length 451;
Best Local Similarity 82.5%; Pred. No. 1.2e-31;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 PLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 61
Db 291 PLPMQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 350
QY 62 NNNQTLFLEHSLLTANTTKG 81
Db 351 NNNQTLSPHSPQMVNTTEG 370

US-09-197-970B-3
; Sequence 3, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; APPLICANT: Joseph V. Bonventre
; APPLICANT: Catherine A. Hession
; APPLICANT: Takaharu Ichimura
; APPLICANT: Henry Wei
; APPLICANT: Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

```
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09941
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/287,001
; FILING DATE: 5 AUG 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09941-2

Query Match 80.6%; Score 348; DB 5; Length 451;
Best Local Similarity 82.5%; Pred. No. 1.2e-31;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 PLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 61
Db 291 PLPMQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDTVTSSDGLW 350
QY 62 NNNQTLFLEHSLLTANTTKG 81
Db 351 NNNQTLSPHSPQMVNTTEG 370

US-09-197-970B-3
; Sequence 3, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; APPLICANT: Joseph V. Bonventre
; APPLICANT: Catherine A. Hession
; APPLICANT: Takaharu Ichimura
; APPLICANT: Henry Wei
; APPLICANT: Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```



```

; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-May-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 307 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-197-970B-3

Query Match 20.8%; Score 89; DB 4; Length 307;
Best Local Similarity 34.1%; Pred. No. 0.036;
Matches 31; Conservative 9; Mismatches 27; Indels 24; Gaps 6;

QY 5 RQNHPEVAT--SPSSPQPAET-----HPTTLOGAIRREPTSSPLYSTTGDND 50
Db 155 RSTHVTSTRVSTSTPTPQTQTHKEITTFYAHETT-----AEVTETPSYT-PADWNG 207

QY 51 TVTESDGLWNNNQTLFLEHSLTANTTKG 81
Db 208 TVT-SSEAWNNHTVPLRKP--QRNPTKG 235

RESULT 5
US-09-006-535-4
; Sequence 4, Application US/09006535
; Patent No. 5965396
; GENERAL INFORMATION:
; APPLICANT: Julie Yan Pan
; APPLICANT: Mark Egerton
; APPLICANT: David Shay Silberstein
; TITLE OF INVENTION: HUMAN LYPH NODE DERIVED GTPase
; FILE REFERENCE: PHM 70295
; CURRENT APPLICATION NUMBER: US/09/006,535
; PRIOR FILING DATE: 1998-01-13
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-006-535-4

Query Match 16.6%; Score 71.5; DB 2; Length 297;
Best Local Similarity 30.6%; Pred. No. 3.5;
Matches 19; Conservative 9; Mismatches 27; Indels 7; Gaps 3;

QY 1 MPLPRQNHPEP--VATSPSPQPAETHTTLOGAIRREPTSSPLYSTTGDNDTVTSSDG 58
Db 21 LPLSSRGHPGRICLTAPSAP---SQHPRLGQSVSLNPPVRKP--SPAQDGSSESSDSG 75

QY 59 LW 60
Db 76 SW 77

RESULT 6
US-09-854-856-64
; Sequence 64, Application US/09854856
```

```

; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 64
; LENGTH: 1911
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1911)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-64

Query Match 16.4%; Score 71; DB 4; Length 1911;
Best Local Similarity 39.6%; Pred. No. 46;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHTTLOGAIRREPTSSPLYSTTGDNDTVTSS 56
Db 715 EPVAVA---QPQATQPTTLASSV--DSAHSDVASGMSDGNENVPSS 756

RESULT 7
US-09-854-856-48
; Sequence 48, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 1939
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1939)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-48

Query Match 16.4%; Score 71; DB 4; Length 1939;
Best Local Similarity 39.6%; Pred. No. 47;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHTTLOGAIRREPTSSPLYSTTGDNDTVTSS 56
Db 743 EPVAVA---QPQATQPTTLASSV--DSAHSDVASGMSDGNENVPSS 784

RESULT 8
US-09-854-856-32
```

```
; Sequence 32, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 1971
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1971)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-32
```

```
Query Match 16.4%; Score 71; DB 4; Length 1971;
Best Local Similarity 39.6%; Pred. No. 48;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESS 56
||||| : ||||| : : : : : : : : : : : :
Db 775 EPVAVA---QPQATQPTTLASSV--DSAHSVDVAGMSDGENVPSS 816
```

```
RESULT 9
US-09-854-856-16
; Sequence 16, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1999
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1999)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-16
```

```
Query Match 16.4%; Score 71; DB 4; Length 1999;
Best Local Similarity 39.6%; Pred. No. 49;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESS 56
||||| : ||||| : : : : : : : : : : : :
Db 803 EPVAVA---QPQATQPTTLASSV--DSAHSVDVAGMSDGENVPSS 844
```

RESULT 10

```
US-09-854-856-58
; Sequence 58, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 58
; LENGTH: 2004
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2004)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-58
```

```
Query Match 16.4%; Score 71; DB 4; Length 2004;
Best Local Similarity 39.6%; Pred. No. 49;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESS 56
||||| : ||||| : : : : : : : : : : : :
Db 808 EPVAVA---QPQATQPTTLASSV--DSAHSVDVAGMSDGENVPSS 849
```

```
RESULT 11
US-09-854-856-42
; Sequence 42, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. 6541252el Human Kinases and Polynucleotides
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; CURRENT FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: US 60/206,015
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42
; LENGTH: 2032
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2032)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-42
```

```
Query Match 16.4%; Score 71; DB 4; Length 2032;
Best Local Similarity 39.6%; Pred. No. 50;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESS 56
||||| : ||||| : : : : : : : : : : : :
Db 836 EPVAVA---QPQATQPTTLASSV--DSAHSVDVAGMSDGENVPSS 877
```

Query	Subject	Score	DB	Length	Mismatches	Indels	Gaps
1	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
2	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
3	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
4	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
5	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
6	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
7	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
8	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
9	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
10	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
11	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
12	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
13	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
14	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
15	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
16	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
17	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
18	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
19	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
20	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
21	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
22	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
23	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
24	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
25	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
26	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
27	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
28	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
29	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
30	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
31	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
32	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
33	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
34	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
35	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
36	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
37	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
38	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
39	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
40	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
41	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
42	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
43	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
44	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
45	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
46	US-09-854-856-46	16.4%	DB 4	2076	7	16	6
47	US-09-854-						

RESULT 12
US-09-854-856-62
; Sequence 62, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 62
; LENGTH: 2048
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2048)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-62

Query Match 16.4%; Score 71; DB 4; Length 2048;
Best Local Similarity 39.6%; Pred. No. 51;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGDNDVTSS 56
||||| : ||||| : : : : : : : : : :
Db 715 EPVAVA----QPQTPTTLASSV--DSAHSVDVASCMSDGNENVPSS 756

RESULT 13
US-09-854-856-26
; Sequence 26, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 2064
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2064)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-26

Query Match 16.4%; Score 71; DB 4; Length 2064;
Best Local Similarity 39.6%; Pred. No. 51;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGDNDVTSS 56
||||| : ||||| : : : : : : : : : :
Db 868 EPVAVA----QPQTPTTLASSV--DSAHSVDVASCMSDGNENVPSS 909

RESULT 14
US-09-854-856-46
; Sequence 46, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 2076
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2076)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-46

Query Match 16.4%; Score 71; DB 4; Length 2076;
Best Local Similarity 39.6%; Pred. No. 51;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGDNDVTSS 56
||||| : ||||| : : : : : : : : : :
Db 743 EPVAVA----QPQTPTTLASSV--DSAHSVDVASCMSDGNENVPSS 784

RESULT 15
US-09-854-856-10
; Sequence 10, Application US/09854856
; Patent No. 6541252
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: Encoding the Same
; FILE REFERENCE: LEX-0178-USA
; CURRENT APPLICATION NUMBER: US/09/854,856
; PRIOR FILING DATE: 2001-05-14
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 2092
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(2092)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-854-856-10

Query Match 16.4%; Score 71; DB 4; Length 2092;
Best Local Similarity 39.6%; Pred. No. 52;
Matches 19; Conservative 7; Mismatches 16; Indels 6; Gaps 2;

QY 9 EPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGDNDVTSS 56
||||| : ||||| : : : : : : : : : :
Db 896 EPVAVA----QPQTPTTLASSV--DSAHSVDVASCMSDGNENVPSS 937

Search completed: June 29, 2005, 09:02:56
Job time : 12.1395 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:46:27 ; Search time 31.4656 Seconds
(without alignments)

635.503 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300

Perfect score: 275

Sequence: 1 PLYSVTTQNDVTWTSDDL.....SLLTANTTKGIYAGVCISVL 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 487488

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	99	36.0	18	17	US-10-718-321-1	Sequence 1, Appl
2	48	17.5	38	16	US-10-425-115-199618	Sequence 199618,
3	48	17.5	44	15	US-10-424-599-190016	Sequence 190016,
4	48	17.5	48	15	US-10-424-599-282496	Sequence 282496,
5	47.5	17.3	40	15	US-10-424-599-264613	Sequence 264613,
6	47	17.1	44	14	US-10-321-857-22	Sequence 22, Appl
7	47	17.1	44	14	US-10-318-675-22	Sequence 22, Appl
8	46	16.7	44	15	US-10-424-599-240616	Sequence 240616,
9	46	16.7	49	15	US-10-424-599-179504	Sequence 179504,
10	45	16.4	37	10	US-09-983-802-399	Sequence 399, App
11	45	16.4	37	10	US-09-984-490-399	Sequence 399, App

12	45	16.4	37	11	US-09-973-278-509	Sequence 509, App
13	43.5	15.8	31	15	US-10-424-599-284397	Sequence 284397,
14	43.5	15.8	42	16	US-10-437-963-166786	Sequence 166786,
15	43.5	15.8	44	16	US-10-437-963-195017	Sequence 195017,
16	43.5	15.8	48	15	US-10-424-599-151333	Sequence 151333,
17	43	15.6	43	9	US-09-925-300-1031	Sequence 1031, Ap
18	42.5	15.5	43	16	US-10-425-115-191318	Sequence 191318,
19	42.5	15.5	45	15	US-10-424-599-152970	Sequence 152970,
20	42.5	15.5	47	16	US-10-425-115-338727	Sequence 338727,
21	42	15.3	35	15	US-10-424-599-216152	Sequence 216152,
22	42	15.3	42	15	US-10-424-599-265756	Sequence 265756,
23	42	15.3	42	16	US-10-425-115-335582	Sequence 335582,
24	42	15.3	47	15	US-10-424-599-277738	Sequence 277738,
25	42	15.3	50	16	US-10-425-115-367358	Sequence 367358,
26	41.5	15.1	37	15	US-10-424-599-269343	Sequence 269343,
27	41.5	15.1	42	15	US-10-424-599-274203	Sequence 274203,
28	41.5	15.1	43	14	US-10-321-857-117	Sequence 117, App
29	41.5	15.1	43	14	US-10-318-675-117	Sequence 117, App
30	41	14.9	30	10	US-09-974-879-549	Sequence 549, App
31	41	14.9	30	10	US-09-305-736-554	Sequence 554, App
32	41	14.9	30	10	US-09-818-683-554	Sequence 554, App
33	41	14.9	30	11	US-09-818-683-554	Sequence 554, App
34	41	14.9	30	15	US-10-621-401-549	Sequence 549, App
35	41	14.9	32	16	US-10-425-115-328576	Sequence 328576,
36	41	14.9	37	16	US-10-425-115-330334	Sequence 330334,
37	41	14.9	39	9	US-09-984-245-123	Sequence 123, App
38	41	14.9	39	10	US-09-966-262-123	Sequence 123, App
39	41	14.9	39	10	US-09-983-966-123	Sequence 123, App
40	41	14.9	39	14	US-10-059-395-123	Sequence 123, App
41	41	14.9	39	14	US-10-143-090-123	Sequence 123, App
42	41	14.9	39	16	US-10-425-115-194419	Sequence 194419,
43	41	14.9	39	16	US-10-608-029-29	Sequence 29, Appl
44	41	14.9	39	17	US-10-960-251-123	Sequence 123, App
45	41	14.9	47	15	US-10-424-599-263904	Sequence 263904,

ALIGNMENTS

RESULT 1

US-10-718-321-1
; Sequence 1, Application US/10718321
; Publication NO. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; TITLE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-1

Query Match 36.0%; Score 99; DB 17; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.1e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 SSDGLWNNQTLFLEHS 33
|||||


```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 44
; TYPE: PR1
; ORGANISM: Lactococcus bacteriophage Tuc2009
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)..(44)
; OTHER INFORMATION: Acma cell wall binding domain homologue
US-10-321-857-22

```

Query Match 17.1%; Score 47; DB 14; Length 44;
Best Local Similarity 32.6%; Pred. No. 1.1e+02;
Matches 15; Conservative 6; Mismatches 21; Indels

Qy 5 YTTDGN^{DVT}ESSDGLWNNQTQLFLEHSLLTANTTKGIYAGVCIS 50
| | | : : | | : : ||||| ||
Db 1 YVVKQGDTLSCIASN-WGNTNWQELARQNLSNPNM--IYAGQVIS 42

```

RESULT 7
US-10-318-675-22
; Sequence 22, Application US/10318675
; Publication No. US20030186851A1
; GENERAL INFORMATION:
; APPLICANT: Applied NanoSystems B.V.
; TITLE OF INVENTION: Bacterial ghosts provided with antigens
; FILE REFERENCE: 2183-5546US
; CURRENT APPLICATION NUMBER: US/10/318,675
; CURRENT FILING DATE: 2002-12-13
; PRIOR APPLICATION NUMBER: PCT/NL02/00383
; PRIOR FILING DATE: 2002-06-11
; PRIOR APPLICATION NUMBER: 01202239.8
; PRIOR FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: Patent In Ver. 2.11

```

```

/ SEQ_ID NO 22
/ LENGTH: 44
/ TYPE: PRT
/ ORGANISM: Lactococcus bacteriophage Tuc2009
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (1)..(44)
/ OTHER INFORMATION: Acma cell wall binding domain homologue
US-0-318-675-22

```

Query Match 17.1%; Score 47; DB 14; Length 44;
Best Local Similarity 32.6%; Pred. No. 1.1e+02;
Matches 15; Conservative 6; Mismatches 21; Indels

Qy 5 YTTDGDNTVTYESSDGLWNNNQTLFLEHSLLTANTTKGIYAGVCIS 50
| | | | : | | : | : | | |
Dd 1 YVVKGGDTLSGIA SN - WGTNWOELARONSLSNP NM -- IYAGQVIS 42

```

RESULT 8
US-10-424-599-240616
; Sequence 240616, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Mol
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Impro
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 240616

```

```

LENGTH: 44
TYPE: PRT

```

```

; ORGANISM: Glycine max
;
; FEATURE:
;
; OTHER INFORMATION: Clone ID: PAT_MRT3847_59303C.1 pep
US-10-424-599-240616

```

Query Match	16.7%	Score 46;	DB 15;
Best Local Similarity	36.8%;	Pred. No. 1.5e+02;	Length 44;
Matches 14;	Conservative	5;	Mismatches 9;
			Indels 10;
			Gaps 3;

[illegible]

RESULT 9
US-10-424-599-179504
; Sequence 179504, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

```

, FILE REFERENCE: 38-211332237B
, CURRENT APPLICATION NUMBER: US/10/424,599
, CURRENT FILING DATE: 2003-04-28
, NUMBER OF SEQ ID NOS: 285684
, SEQ ID NO 179504
, LENGTH: 49
, TYPE: PRT
, ORGANISM: Glycine max
, FEATURE:
, OTHER INFORMATION: Clone ID: PAT_MRT3847_133108C.1 pep
US-10-424-599-179504

```

Query Match	16.7%	Score 46;	DB 15;	Length 49;
Best Local Similarity	27.3%;	Pred. No. 1.7e+02;		
Matches 12:	Conservative	8;	Mismatches 22;	Indels 2;
				Gaps 1;

Qy	10	NDVTTESSDGLWNNNQTLPL--EHSLLTANTTKGIYAGVCISV	51
		: :	
Db	2	SDNAKSSADSYINDRSVDYLIDYEHLVSSNTFKNYTRHGCDKV	45

RESULT	10
US-09-983-802-399	
Sequence 399, Application US/09983802	
Publication No. US20030022185A1	
GENERAL INFORMATION:	
APPLICANT: Fischer et al.	
TITLE OF INVENTION: 123 Human Secreted Proteins	
FILE REFERENCE: PZ01OP1	
CURRENT APPLICATION NUMBER: US/09/983,802	
CURRENT FILING DATE: 2001-10-25	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/227,357	
PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-08	PCT/US98/13
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/051,926
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/052,793
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/051,925
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/051,929
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/052,803
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/052,732
PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08	
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:	60/051,931

```
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,932
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,722
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,954
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,785
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,664
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,660
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,661
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 399
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-983-802-399
```

```
Query Match 16.4%; Score 45; DB 10; Length 37;
Best Local Similarity 27.3%; Pred. No. 1.7e+02;
Matches 9; Conservative 7; Mismatches 13; Indels 4; Gaps 1;
```

```
QY 20 LNNNQTLFLBLSLLTANTTGGVAGVCISVL 52
||| | | | | | | | | | | | | | | | |
Db 3 LNNRNQ----NMHSIIVKELIVTFGLGITVLL 31
```

```
RESULT 11
US-09-984-490-399
; Sequence 399, Application US/09984490
; Publication No. US2003006412A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P1
; CURRENT APPLICATION NUMBER: US/09/984,490
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/227,357
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: PCT/US98/13684
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,926
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,793
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,925
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,929
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,803
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,732
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,931
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,932
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,916
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,930
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,918
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,920
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,733
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/052,795
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,722
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,723
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,948
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,949
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,953
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,950
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,947
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,964
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/056,360
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/055,954
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,785
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,664
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,660
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/058,661
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 399
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-983-802-399
```


; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 399
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-490-399

Query Match 16.4%; Score 45; DB 10; Length 37;
Best Local Similarity 27.3%; Pred. No. 1.7e+02;
Matches 9; Conservative 7; Mismatches 13; Indels 4; Gaps 1;

QY 20 LWNNTQQLFLEHSLTANTTKGIYAGVCISVL 52
||| : : : : :
DB 3 LWNRNQ----MMHSIIIVKELIVTFELGITVLL 31

RESULT 12

US-09-973-278-509
; Sequence 509, Application US/09973278
; Publication No. US20040044191A1
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: PZ010P2
; CURRENT APPLICATION NUMBER: US/09/973,278
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/239,899
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/227,357
; PRIOR FILING DATE: 1999-01-08
; PRIOR APPLICATION NUMBER: PCT/US98/13684
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/051,926
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,793
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,925
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,929
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,803
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,732
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,931
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,932
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,916
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,930
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,918
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,920
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,733
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/052,795
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,919
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/051,928
; PRIOR FILING DATE: 1997-07-08
; PRIOR APPLICATION NUMBER: 60/055,722
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,723
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,948
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,949
; PRIOR FILING DATE: 1997-08-18

; PRIOR APPLICATION NUMBER: 60/055,953
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,950
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,947
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,964
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/056,360
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,684
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,984
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/055,954
; PRIOR FILING DATE: 1997-08-18
; PRIOR APPLICATION NUMBER: 60/058,785
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,664
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,660
; PRIOR FILING DATE: 1997-09-12
; PRIOR APPLICATION NUMBER: 60/058,661
; PRIOR FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 947
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 509
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-278-509

Query Match 16.4%; Score 45; DB 11; Length 37;
Best Local Similarity 27.3%; Pred. No. 1.7e+02;
Matches 9; Conservative 7; Mismatches 13; Indels 4; Gaps 1;

QY 20 LWNNTQQLFLEHSLTANTTKGIYAGVCISVL 52
||| : : : : :
DB 3 LWNRNQ----MMHSIIIVKELIVTFELGITVLL 31

RESULT 13

US-10-424-599-284397
; Sequence 284397, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 284397
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98836C.1.pap
US-10-424-599-284397

Query Match 15.8%; Score 43.5; DB 15; Length 31;
Best Local Similarity 37.5%; Pred. No. 2.1e+02;
Matches 9; Conservative 7; Mismatches 5; Indels 3; Gaps 1;

QY 21 WNNNOT---QLFLEHSLTANTTK 41
|:|:| : : : : :
DB 5 WDNSTYLVQFLENQLLSLSASK 28

Job time : 32.4656 secs

```
RESULT 14
US-10-437-963-166786
; Sequence 166786, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 166786
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_65460C.1.pep
US-10-437-963-166786

Query Match      15.8%; Score 43.5; DB 16; Length 42;
Best Local Similarity 47.8%; Pred. No. 3.1e+02;
Matches 11; Conservative 6; Mismatches 3; Indels 3; Gaps 2;

QY      20 LWNNTQTLF-LEHSLTANTTK 41
DB      15 IWTNLTKLYLLKHTLL--NSTK 35

RESULT 15
US-10-437-963-195017
; Sequence 195017, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195017
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91005C.1.pep
US-10-437-963-195017

Query Match      15.8%; Score 43.5; DB 16; Length 44;
Best Local Similarity 35.3%; Pred. No. 3.3e+02;
Matches 12; Conservative 8; Mismatches 11; Indels 3; Gaps 2;

QY      21 WNNNTQLFLEHSL--LTANTTKGIYAGVCISVL 52
DB      3 WHSLPDEVW-EHALSFLPADADRGAAGACTSWL 35

Search completed: June 29, 2005, 10:14:20
```

Result No.	Score	Query Match	Length	DB	ID	Description
1	52.5	19.1	30	4	US-09-282-029A-470	Sequence 470, App
2	52.5	19.1	30	4	US-09-282-029A-471	Sequence 471, App
3	52.5	19.1	30	4	US-09-434-355A-470	Sequence 470, App
4	52.5	19.1	30	4	US-09-434-355A-471	Sequence 471, App
5	45.5	16.5	40	1	US-08-188-228-22	Sequence 22, Appl
6	45.5	16.5	40	1	US-08-332-643-22	Sequence 22, Appl
7	45.5	16.5	40	1	US-08-332-638-22	Sequence 22, Appl
8	45	16.4	37	3	US-09-227-357-399	Sequence 399, App
9	44.5	16.2	31	3	US-08-190-802A-244	Sequence 244, App
10	44.5	16.2	31	3	US-08-477-346-244	Sequence 244, App
11	44.5	16.2	31	3	US-08-477-389-244	Sequence 244, App
12	44.5	16.2	31	3	US-08-487-072A-244	Sequence 244, App
13	44	16.0	40	1	US-08-218-025A-65	Sequence 65, Appl
14	42	15.3	40	1	US-08-188-228-16	Sequence 16, Appl
15	42	15.3	40	1	US-08-332-643-16	Sequence 16, Appl
16	42	15.3	40	1	US-08-332-638-16	Sequence 16, Appl
17	41.5	15.1	32	2	US-08-708-620A-4	Sequence 4, Appl
18	41	14.9	42	4	US-09-434-355A-529	Sequence 529, App
19	40	14.5	19	4	US-09-570-921-141	Sequence 141, App
20	40	14.5	42	4	US-09-282-029A-6	Sequence 6, Appl
21	40	14.5	42	4	US-09-185-908-6	Sequence 6, Appl
22	40	14.5	42	4	US-09-434-355A-6	Sequence 6, Appl
23	40	14.5	49	3	US-09-217-228-9	Sequence 9, Appl
24	39.5	14.4	30	1	US-08-190-802A-123	Sequence 123, App
25	39.5	14.4	30	3	US-08-477-346-123	Sequence 123, App
26	39.5	14.4	30	3	US-08-473-089-123	Sequence 123, App
27	39.5	14.4	30	4	US-08-487-072A-123	Sequence 123, App


```
/
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/332,643
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/872,643
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 5639634and, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 27866/30795
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 346-5750
/ TELEFAX: (312) 984-9740
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 40 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-332-643-22

Query Match      16.5%; Score 45.5; DB 1; Length 40;
Best Local Similarity 31.7%; Pred. No. 45;
Matches 13; Conservative 5; Mismatches 8; Indels 15; Gaps 1;

QY  2 LYSYTTDGNQDVTSSDGLWNNNQTLFLEHLLTANTTKG 42
    | : | : | : | : | : | : | : | : | : | : |
Db   2 LATYAYEGNDSVAESL-----SSLESGTTEG 27

RESULT 7
US-08-332-638-22
/ Sequence 22, Application US/08332638
/ Patent No. 5646250
/ GENERAL INFORMATION:
/ APPLICANT: Suzuki, Shintaro
/ TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
/ NUMBER OF SEQUENCES: 62
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
/ ADDRESS: 6300 Sears Tower, 233 S. Wacker Drive
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/332,638
/ FILING DATE: 01-NOV-1994
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/872,643
/ FILING DATE: 17 APR 1992
/ APPLICATION NUMBER: US/08/049,460
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 5646250and, Greta E.
/ REGISTRATION NUMBER: 35,302
/ REFERENCE/DOCKET NUMBER: 31340
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (312) 474-6300
/ TELEFAX: (312) 474-0448
/ TELEX: 25-3856
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 40 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-332-638-22

Query Match      16.5%; Score 45.5; DB 1; Length 40;
Best Local Similarity 31.7%; Pred. No. 45;
Matches 13; Conservative 5; Mismatches 8; Indels 15; Gaps 1;

QY  2 LYSYTTDGNQDVTSSDGLWNNNQTLFLEHLLTANTTKG 42
    | : | : | : | : | : | : | : | : | : | : |
Db   2 LATYAYEGNDSVAESL-----SSLESGTTEG 27

RESULT 8
US-09-227-357-399
/ Sequence 399, Application US/09227357
/ Patent No. 6342581
/ GENERAL INFORMATION:
/ APPLICANT: Fischer et al.
/ TITLE OF INVENTION: 123 Human Secreted Proteins
/ FILE REFERENCE: P2010P1
/ CURRENT APPLICATION NUMBER: US/09/227,357
/ CURRENT FILING DATE: 1999-01-08
/ EARLIER APPLICATION NUMBER: PCT/US98/13684
/ EARLIER FILING DATE: 1998-07-07
/ EARLIER APPLICATION NUMBER: 60/051,926
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/052,793
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,925
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,929
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/052,803
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/052,732
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,931
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,932
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,916
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,930
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,918
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,920
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/052,733
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/052,795
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,919
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/051,928
/ EARLIER FILING DATE: 1997-07-08
/ EARLIER APPLICATION NUMBER: 60/055,722
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,723
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,948
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,949
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,953
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,950
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,947
/ EARLIER FILING DATE: 1997-08-18
/ EARLIER APPLICATION NUMBER: 60/055,964
```

```
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,684
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,984
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,954
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/058,785
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,664
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,660
; EARLIER FILING DATE: 1997-09-12
; EARLIER APPLICATION NUMBER: 60/058,661
; EARLIER FILING DATE: 1997-09-12
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 399
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-227-357-399

Query Match          16.4%; Score 45; DB 3; Length 37;
Best Local Similarity 27.3%; Pred. No. 47;
Matches 9; Conservative 7; Mismatches 13; Indels 4; Gaps 1;

QY 20 LWNNTQTLFLEHSLTANTTKGIYAGVCISVL 52
   ||||| :||| :||| :||| :||| :|||
DB 3 LWNRNQ----MMHSIIVKELIVTFGLITVLL 31

RESULT 9
US-08-190-802A-244
; Sequence 244, Application US/08190802A
; Patent No. 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Derlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
```

```
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: YCU7 rI, Fig. 48
US-08-190-802A-244

Query Match          16.2%; Score 44.5; DB 1; Length 31;
Best Local Similarity 40.9%; Pred. No. 44;
Matches 9; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 4 SYTTDNDTVTSSDG---LWN 22
   :|: ||: ||| ||| ||| :|:
DB 10 AYSPDGSRVVTASEDGKIKWVD 31

RESULT 10
US-08-477-346-244
; Sequence 244, Application US/08477346
; Patent No. 6262023
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08477,346
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,072
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MORASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: YCU7 rI, Fig. 48
US-08-477-346-244

Query Match          16.2%; Score 44.5; DB 3; Length 31;
Best Local Similarity 40.9%; Pred. No. 44;
Matches 9; Conservative 5; Mismatches 5; Indels 3; Gaps 1;

QY 4 SYTTDNDTVTSSDG---LWN 22
   :|: ||: ||| ||| ||| :|:
DB 10 AYSPDGSRVVTASEDGKIKWVD 31

RESULT 11
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,072A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 244:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: YCU7 rI, Fig. 48
; US-08-487-072A-244

Query Match 16.2%; Score 44.5; DB 4; Length 31;
Best Local Similarity 40.9%; Pred. No. 44;
Matches 9; Conservative 5; Mismatches 5; Indels 3;

QY      4 SYTTDGNDTVTSSDG---LWN 22
        :|: ||: || | || | |: 
DB       10 AYSPPGSRVVTASEDGKIKWVD 31

RESULT 13
US-08-218-025A-65
; Sequence 65, Application US/08218025A
; Patent No. 5556744
; GENERAL INFORMATION:
; APPLICANT: Weiner, David B.
; APPLICANT: Ugen, Kenneth E.
; APPLICANT: Williams, William V.
; TITLE OF INVENTION: Methods and Compositions for Diagnosing
; TITLE OF INVENTION: and Treating Certain HIV Infected Patients
; NUMBER OF SEQUENCES: 197
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: P.O. Box 457, 321 No. 5556744--rilstown Road
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/218,025A
; FILING DATE: 24-MAR-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/891,451
; FILING DATE: 29-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST33A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 540-9206
; TELEFAX: (215) 540-5818
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:

```

```
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-218-025A-65

; Query Match      16.0%; Score 44; DB 1; Length 20;
; Best Local Similarity 39.1%; Pred. No. 29;
Matches 9; Conservative 4; Mismatches 2; Indels 8; Gaps 1;

QY 1 LYSYTTDGDNTVTSSDGLWNN 23
   ||:|:|:|:|:|
Db 2 PLFNSTWNGNNT-----WNN 16

RESULT 14
US-08-188-228-16
; Sequence 16, Application US/08188228
; Patent No. 5597725
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 S. Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/188,228
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/049,460
; FILING DATE: 19 APR 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/872,643
; FILING DATE: 17 APR 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5597725and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 31340
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-188-228-16

Query Match      15.3%; Score 42; DB 1; Length 40;
; Best Local Similarity 39.1%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 LYSYTTDGDNTVTSSDGLWNN 24
   ||:|:|:|:|:|
Db 2 LQYAFEGNGSVAESLSLDSNS 24

RESULT 15
US-08-332-643-16
```

```
; Sequence 16, Application US/08332643
; Patent No. 5639634
; GENERAL INFORMATION:
; APPLICANT: Suzuki, Shintaro
; TITLE OF INVENTION: CADHERIN MATERIALS AND METHODS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Bicknell
; STREET: Two first National Plaza, 20 South Clark
; STREET: Street
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/332,643
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/872,643
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5639634and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/30795
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 346-5750
; TELEFAX: (312) 984-9740
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-332-643-16

Query Match      15.3%; Score 42; DB 1; Length 40;
; Best Local Similarity 39.1%; Pred. No. 1.3e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 LYSYTTDGDNTVTSSDGLWNN 24
   ||:|:|:|:|:|
Db 2 LQYAFEGNGSVAESLSLDSNS 24

Search completed: June 29, 2005, 09:49:08
Job time : 9.39882 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 09:03:05 ; Search time 35.5521 Seconds
(without alignments)
565.693 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300
Perfect score: 275
Sequence: 1 PLYSVTTGNDVTFTSSDGL.....SLLTANTTKGIYAGVICISVL 52

Scoring table: BLASTSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 385760381 residues

Total number of hits satisfying chosen parameters: 938430

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	101	36.7	18	6	Aao26686 Human KIM
2	99	36.0	18	6	Aao26687 Human KIM
3	99	36.0	18	6	Aao26673 Monoclonal
4	90	32.7	18	6	Aao26688 Human KIM
5	82	29.8	18	6	Aao26685 Human KIM
6	49	17.8	43	6	Ada98306 Human sec
7	49	17.8	43	6	Ada44132 Human sec
8	49	17.8	43	7	Adc20470 Human sec
9	49	17.8	43	7	Adf10760 Human sec
10	49	17.8	44	3	Aab29810 Human sec
11	48.5	17.6	40	8	Ad807403 Staphyloc
12	47	17.1	44	2	Aay13969 Lys repa
13	47	17.1	44	6	Abr83729 Bacteriop
14	47	17.1	44	7	Adf70020 Acma-type
15	46	16.7	49	4	Aam92081 Human dig
16	45.5	16.5	40	2	Aaw25651 Rat cadhe
17	45.5	16.5	40	2	Aaw13116 Rat cadhe
18	45	16.4	37	7	Ada07578 Human sec
19	45	16.4	37	8	Adn41387 Novel hum
20	44.5	16.2	31	2	Aar85818 Peptide r
21	44	16.0	20	2	Aaw08045 HIV pepti
22	44	16.0	44	7	Adf70018 Acma-type
23	43	15.6	43	3	Aab56453 Human pro
24	42.5	15.5	20	8	Ad195222 OSPF-rela
25	42	15.3	40	2	Aar43565 Rat cadhe

26	42	15.3	40	2	AAW25648	Aaw25648 Rat cadhe
27	42	15.3	40	2	AAW13113	Aaw13113 Rat cadhe
28	42	15.3	46	7	ADF70089	Adf70089 Acma-type
29	42	15.3	49	2	AAR83059	Aar83059 Human IL-
30	41.5	15.1	32	2	AAW18789	Aaw18789 Z peptide
31	41.5	15.1	43	6	ABR83824	Abr83824 Bacillus
32	41.5	15.1	43	7	ADF69876	Adf69876 Acma-type
33	41.5	15.1	49	2	AAR61126	Aar61126 Vasoactiv
34	41	14.9	30	6	ABO14381	Abol14381 Novel hum
35	41	14.9	30	8	ADG78793	Adg78793 Human sec
36	41	14.9	30	8	ADN61079	Adn61079 Human sec
37	41	14.9	39	2	AAW67809	Aaw67809 Human sec
38	41	14.9	46	4	AAW24315	Aam24315 Rat EST e
39	41	14.9	46	7	ADF60572	Adf60572 Human con
40	41	14.9	46	7	ADF59574	Adf59574 Human pol
41	41	14.9	49	1	AAP91538	Aap91538 Modified
42	41	14.9	50	5	ABP11026	Abp11026 Human ORF
43	40.5	14.7	50	4	ABG30194	Abg30194 Novel hum
44	40	14.5	14	7	ADE03490	Ade03490 BGS-3 asp
45	40	14.5	14	8	ADO10179	Ado10179 Human int

ALIGNMENTS

RESULT 1
AAO26686
ID AAO26686 standard; peptide; 18 AA.
XX
AC AAO26686;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain related peptide, SEQ ID No 14.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
OS Synthetic.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
(BIOJ) BIOGEN INC.
(GEHO) GEN HOSPITAL CORP.
PI Bailly V, Bonventre J;
XX
DR WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
e.g. renal cancer.
XX
PS Disclosure; Fig 1A; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
XX peptide of the invention

```

SQ Sequence 18 AA;
  Query Match          36.7%; Score 101; DB 6; Length 18;
  Best Local Similarity 100.0%; Pred. No. 1.4e-05;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7 TDGNDTVTESSDGLWNNN 24
    |||||
Db 1 TDGNDTVTESSDGLWNNN 18
    |||||

RESULT 2
AAO26687
ID AAO26687 standard; peptide; 18 AA.
AC AAO26687;
XX
XX 20-MAR-2003 (first entry)
XX Human KIM-1 mucin domain related peptide, SEQ ID No 15.
DE
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
XX Synthetic.
XX
XX WO200298920-A1.
XX
XX 12-DEC-2002.
XX
XX 31-MAY-2002; 2002WO-US017402.
XX
XX 01-JUN-2001; 2001US-0295449P.
XX 04-JUN-2001; 2001US-0295907P.
XX
XX (BIOJ ) BIOGEN INC.
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Bailly V, Bonventre J;
XX
XX WPI; 2003-156845/15.
XX
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Claim 4; Page 25; 42pp; English.
XX
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain related
CC peptide of the invention
XX
XX Sequence 18 AA;
  Query Match          36.0%; Score 99; DB 6; Length 18;
  Best Local Similarity 100.0%; Pred. No. 2.5e-05;
  Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 SSDGLWNNNTQLFLEHS 33
    |||||
Db 1 SSDGLWNNNTQLFLEHS 18
    |||||

RESULT 4
AAO26688
ID AAO26688 standard; peptide; 18 AA.
XX
XX AAO26688;
XX
XX 20-MAR-2003 (first entry)
XX
XX Human KIM-1 mucin domain related peptide, SEQ ID No 16.
DE
XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer.
XX
XX Synthetic.
XX
XX WO200298920-A1.
XX

```

PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-US017402.
 XX
 PR 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (GHEO) GEN HOSPITAL CORP.
 XX
 PI Bailly V, Bonventre J;
 XX
 DR WPI; 2003-156845/15.
 XX
 PT New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Disclosure; Fig 1A; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX
 SQ Sequence 18 AA;
 Query Match 32.7%; Score 90; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 25 QTQLFLEHSLTANTTKG 42
 DB 1 QTQLFLEHSLTANTTKG 18
 RESULT 5
 ID AAO26685 standard; peptide; 18 AA.
 XX
 AC AAO26685;
 XX
 DT 20-MAR-2003 (first entry)
 XX
 DE Human KIM-1 mucin domain related peptide, SEQ ID No 13.
 XX
 KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
 KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
 KW renal disease; injury; renal cancer.
 XX
 OS Synthetic.
 XX
 PN WO200298920-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 31-MAY-2002; 2002WO-US017402.
 XX
 PR 01-JUN-2001; 2001US-0295449P.
 PR 04-JUN-2001; 2001US-0295907P.
 XX
 PA (BIOJ) BIOGEN INC.
 PA (GHEO) GEN HOSPITAL CORP.
 XX
 PI Bailly V, Bonventre J;
 XX
 DR WPI; 2003-156845/15.
 XX

PT New antibody, antibody derivative or antigen-binding polypeptide that
 PT inhibits proteolytic release of a soluble kidney injury molecule-1
 PT polypeptide, useful for treating or preventing renal disease or injury,
 PT e.g. renal cancer.
 XX
 PS Disclosure; Fig 1A; 42pp; English.
 XX
 CC The invention relates to a novel antibody, antibody derivative or antigen
 CC -binding polypeptide that inhibits proteolytic release of a soluble
 CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
 CC The antibody, antibody derivative or antigen-binding polypeptide is
 CC useful for treating or preventing renal disease or injury, e.g. renal
 CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
 CC polypeptide. This sequence represents a human KIM-1 mucin domain related
 CC peptide of the invention
 XX
 SQ Sequence 18 AA;
 Query Match 29.8%; Score 82; DB 6; Length 18;
 Best Local Similarity 100.0%; Pred. No. 0.0034;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PLYSYTTDGNDRVTE 15
 DB 4 PLYSYTTDGNDRVTE 18
 RESULT 6
 ID ADA98306 standard; protein; 43 AA.
 XX
 AC ADA98306;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Human secreted protein sequence #147.
 XX
 KW human; secreted protein; cardiovascular disorder; arrhythmia;
 KW atherosclerosis; stroke; endocarditis; congestive heart failure;
 KW rheumatic heart disease; cardiomyopathy; haemorrhoids; varicose veins;
 KW migraine; thrombosis; neural disorder; immune system disorder;
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;
 KW pulmonary disorder; renal disorder; proliferative disorder; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO2003004623-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 26-MAR-2002; 2002WO-US009922.
 XX
 PR 27-MAR-2001; 2001US-0278650P.
 PR 12-SEP-2001; 2001US-00950082.
 PR 12-SEP-2001; 2001US-00950083.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-247946/24.
 XX
 PT New human secreted polypeptide and nucleic acid molecules, useful for
 PT diagnosing, preventing, prognosticating or treating cardiovascular
 PT disorders (e.g. arrhythmia, atherosclerosis, cardiomyopathy, or
 PT thrombosis).
 XX
 PS Claim 1; SEQ ID NO 414; 1572pp; English.
 XX
 CC The invention comprises the amino acid and coding sequence of human
 CC secreted proteins. The DNA and protein sequences of the invention are
 CC useful in the treatment of cardiovascular disorders, such as: arrhythmia,
 CC atherosclerosis, stroke, endocarditis, congestive heart failure,


```
Db 3 LMKCGRTT--LRASLLMIFTSCSWMLGVCLSVL 33
RESULT 9
ADP10760
ID ADF10760 standard; protein; 43 AA.
XX
AC ADF10760;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human secreted protein #82.
XX
KW H6EDM64; HBHA05; HB3CR46; HB3KDL6; HCWSX51; HCQBH72; HDPQ30; HE2CM39;
KW HE9EA10; HGBHP91; HLDQU79; Cytostatic; Hepatotropic; Antidiabetic;
KW Antinflammatory; neuroprotective; Anti-HIV; Vulnery; Gynecological;
KW Antinfertility; Gene therapy; Gastrointestinal disorder; cancer;
KW Alzheimer's disease; chromosome identification.
XX
OS Homo sapiens.
XX
PN WO200299085-A2.
XX
PD 12-DEC-2002.
XX
PF 26-MAR-2002; 2002WO-US009135.
XX
PR 27-MAR-2001; 2001US-0278650P.
XX
PR 12-SEP-2001; 2001US-00950082.
XX
PR 12-SEP-2001; 2001US-00950083.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI; 2003-221310/21.
XX
PT New human secreted polypeptides for diagnosing and treating neural,
PT immune system, muscular, reproductive, gastrointestinal, cardiovascular,
PT renal, and proliferative disorders and cancerous diseases.
XX
PS Claim 1; SEQ ID NO 223; 855pp; English.
XX
CC The present invention relates to an isolated polypeptide chosen from 123
CC human secreted proteins, such as, H6EDM64, HBHA05, HB3CR46, HB3KDL6,
CC HCWSX51, HCQBH72, HDPQ30, HE2CM39, HE9EA10, HGBHP91 and HLDQU79. The
CC polypeptides are useful for the preparation of a diagnostic or
CC pharmaceutical composition for diagnosing or and are useful for treating
CC or preventing diseases or conditions, such as neural, immune system,
CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
CC renal, proliferative disorders and cancerous diseases and conditions. The
CC polypeptides have immune activity, chemotactic activity, and binding
CC activity to treat and prevent neuronal damage which occurs in certain
CC neuronal disorders or neuro-degenerative conditions such as Alzheimer's
CC disease, Parkinson's disease, and acquired immunodeficiency syndrome
CC (AIDS)-related complex, and to prevent skin aging due to sunburn by
CC stimulating keratinocyte growth. The molecules are also useful to
CC modulate mammalian characteristics including . The encoding sequences are
CC useful for chromosome identification, radiation hybrid mapping, in gene
CC therapy, for identifying individuals from minute biological samples, as
CC additional DNA markers for restriction fragment length polymorphism
CC (RFLP), in forensic biology, molecular weight markers on Southern gels,
CC as diagnostic probes for the presence of a specific mRNA in a particular
CC cell type, to raise anti-DNA antibodies using DNA immunization
CC techniques, and as an antigen to elicit an immune response. The present
CC sequence represents a human secreted protein of the invention.
XX
SQ Sequence 43 AA;
Query Match 17.8%; Score 49; DB 7; Length 43;
Best Local Similarity 42.4%; Pred. No. 1.3e+02;
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

Db 3 LMKCGRTT--LRASLLMIFTSCSWMLGVCLSVL 33
RESULT 10
AAB29810
ID AAB29810 standard; protein; 44 AA.
XX
AC AAB29810;
XX
DT 09-FEB-2001 (first entry)
XX
DE Human secreted protein SEQ ID NO: 68.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW Antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW Vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW Cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
PN WO200061779-A1.
XX
PD 19-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US009068.
XX
PR 09-APR-1999; 99US-0128699P.
XX
PR 20-JAN-2000; 2000US-0177050P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis G;
XX
DR WPI; 2000-647424/62.
XX
DR N-PSDB; AAC63418.
XX
PT Isolated nucleic acid molecule encoding a human secreted protein is used
PT in preventing, treating or ameliorating a medical condition.
XX
PS Claim 11; Page 431; 495pp; English.
XX
CC Sequences AAB29802-B29850 represent the amino acid sequences of 49 human
CC secreted proteins encoded by the genes AAC63410-C63458. The genes and
CC proteins are useful for preventing, ameliorating or treating medical
CC conditions, e.g. by protein or gene therapy. The genes are isolated from
CC a range of human tissues disclosed in the specification. The nucleic
CC acids, proteins, antibodies and (ant)agonists are useful in the
CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections
XX
SQ Sequence 44 AA;
Query Match 17.8%; Score 49; DB 3; Length 44;
Best Local Similarity 42.4%; Pred. No. 1.4e+02;
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

QY 20 LMNNQTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 3 LMKCGRTT--LRASLLMIFTSCSWMLGVCLSVL 33
RESULT 11
QY 20 LMNNQTQLFLEHSLLTANTTKGIYAGVCISVL 52
DB 3 LMKCGRTT--LRASLLMIFTSCSWMLGVCLSVL 33
RESULT 11
```

ADSO7403
ID ADS07403 standard; protein; 40 AA.
XX
AC ADS07403;
XX
DT 04-NOV-2004 (first entry)
XX
DE Staphylococcus epidermis polypeptide seqid 6698.
XX
KW antibacterial; vaccine; antisense therapy; Staphylococcus epidermidis;
KW recombinant expression vector; infection; computer readable medium;
KW computer based system.
XX
OS Staphylococcus epidermidis.
XX
PN US2004147734-A1.
XX
PD 29-JUL-2004.
XX
PF 01-DEC-2003; 2003US-00724972.
XX
PR 08-NOV-1997; 97US-0064964P.
PR 13-AUG-1998; 98US-00134001.
PR 29-NOV-1999; 99US-00450969.
XX
PA (DOUC/) DOUCETTE-STAMM L.
PA (BUSH/) BUSH D.
XX
PI Doucette-Stamm L, Bush D;
XX
DR WPI; 2004-580138/56.
DR N-PSDB; ADS03631.
XX
PT New isolated polypeptide and encoding nucleic acid derived from
PT Staphylococcus epidermidis, useful for diagnosing, preventing and/or
PT treating an S. epidermidis bacterial infection.
XX
PS Claim 17; SEQ ID NO 6698; 741bp; English.
XX
CC The invention describes an isolated nucleic acid comprising a nucleotide
CC sequence with any of 3772 fully defined nucleotide sequences (SEQ ID NO:
CC 1-3772) and encoding an Staphylococcus epidermidis polypeptide with any
CC of 3772 fully defined amino acid sequences (SEQ ID NO: 3772-7544) as
CC given in the specification. Also described are: a recombinant expression
CC vector; a cell comprising a recombinant expression vector of (1);
CC producing an S. epidermidis polypeptide; an isolated nucleic acid
CC comprising a nucleotide sequence of at least 8 nucleotides in length; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection, comprising a nucleic acid cited above and a carrier; treating
CC a subject for S. epidermidis infection; a recombinant or substantially
CC pure preparation of an S. epidermidis polypeptide or its fragment; a
CC vaccine composition for prevention or treatment of an S. epidermidis
CC infection; detecting the presence of a Staphylococcus nucleic acid in a
CC sample; a computer readable medium having recorded in it the nucleotide
CC sequences with SEQ ID NO: 1-3772 or its fragments; a computer based
CC system for identifying fragments of the Staphylococcus genome of
CC commercial importance; a computer based system for identifying fragments
CC of the Staphylococcus plasmids of commercial importance; identifying
CC commercially important nucleic acid fragments of the Staphylococcus
CC genome and/or plasmids; and identifying an expression modulating fragment
CC of the Staphylococcus genome and/or plasmids. The methods and
CC compositions of the present invention are useful for the diagnosis,
CC prevention and/or treatment of an Staphylococcus epidermidis bacterial
CC infection. This is the amino acid sequence of a S. epidermidis protein of
CC the invention.
XX
SQ Sequence 40 AA;
Query Match 17.6%; Score 48.5; DB 8; Length 40;
Best Local Similarity 33.3%; Pred. No. 1.4e+02;
Matches 12; Conservative 9; Mismatches 14; Indels 1; Gaps 1;
QY 2 LYSVTTDGNVTWESSDGLWNNN-QTQLFLEHSLLT 36

DB 1 LYFFTLQONPSFIQCSGVLKNSLQSAIFVLDPLLT 36
||:| | : : |:|: |: |:|: |||
RESULT 12
AAV13969
ID AAV13969 standard; peptide; 44 AA.
XX
AC AAV13969;
XX
DT 15-JUL-1999 (first entry)
XX
DE Lys repeat sequence from Lactococcus.
XX
KW Acma repeat; consensus sequence; major peptidoglycan hydrolase; vaccine;
KW cell wall attachment; substance delivery; diagnosis; bioadsorption.
XX
OS Lactococcus sp.
XX
PN EP916726-A1.
XX
PD 19-MAY-1999.
XX
PF 13-NOV-1997; 97EP-00203539.
XX
PR 13-NOV-1997; 97EP-00203539.
XX
PA (UYGR-) RIJKSUNIV GRONINGEN.
XX
DR WPI; 1999-290024/25.
XX
PT New proteinaceous substance comprising a sequence consensus to a major
PT peptidoglycan (Acma), useful for attaching a substance to a cell wall.
XX
PS Claim 2; Fig 11; 98pp; English.
XX
CC This sequence represents a repeat sequence that can be used in the
CC proteinaceous substance of the invention. The proteinaceous substance
CC comprises at least one stretch of amino acids derived from a first
CC organism, capable of attaching to a cell wall of a second microorganism.
CC The proteinaceous substance is useful in a method for attaching a
CC substance to the cell wall of a microorganism, and the substance and
CC either microorganism are useful in pharmaceutical compositions and
CC vaccines, for delivery of a substance to a cell. They are also useful in
CC diagnostic tests, bioadsorption processes and in foodstuffs. The new
CC method targets substances to cells of a wide range of microorganisms,
CC unlike prior art anchoring and targeting proteins which are specific and
CC selective for a limited set of microorganisms, which are usually
CC recombinant or pathogenic. The second microorganism in the new method is
CC non-recombinant, preventing restrictions on applications, and preventing
CC potential problems of colonisation of the mucosal surfaces which
CC generates long term exposure to the target antigens expressed, which can
CC cause immune tolerance. Public consensus is against use of recombinant or
CC attenuated strains, so the new technique is more likely to be accepted
CC than prior art methods
XX
SQ Sequence 44 AA;
Query Match 17.1%; Score 47; DB 2; Length 44;
Best Local Similarity 32.6%; Pred. No. 2.5e+02;
Matches 15; Conservative 6; Mismatches 21; Indels 4; Gaps 2;
QY 5 YTTDGNVTWESSDGLWNNNQTQLFLEHSLLTANTTKGIYAGVCIS 50
||:| | : : |:|: |: |:|: |||
DB 1 YVVKQDGLTSLGIASN-WGTNWOELARONSLSNPNM---IYAGQVIS 42
RESULT 13
ABR83729
ID ABR83729 standard; peptide; 44 AA.
XX
AC ABR83729;
XX

DT 15-OCT-2003 (first entry)
XX Bacteriophage Tuc2009 lys protein amino acids 332-375.
DE Antibacterial; protozoacide; vaccine; cell-wall material; anchor protein;
XX Gram-positive bacterium; Lactococcus lactis; Acma; biocatalyst;
KW mucosal immunization; bioadsorbent; biofilter; microbiocatalyst;
KW vaccination; malaria; Streptococcus pneumoniae; fusion gene;
KW Streptococcus pneumoniae; protease maturation protein.
XX Bacteriophage Tuc2009.
XX WO2002101026-A2.
FN 19-DEC-2002.
XX 11-JUN-2002; 2002WO-NL000383.
PF 11-JUN-2001; 2001EP-00202239.
XX (NANO-) APPLIED NANOSYSTEMS BV.
PA Leenhouts CJ, Ramasamy R, Steen A, Kok J, Buist G, Kuipers OP;
XX WPI; 2003-167404/16.
DR Improving binding of a proteinaceous substance e.g. an AcmaA-type protein
XX to a cell-wall material of microorganisms, comprises treating the
PT material with a solution capable of removing protein or carbohydrate from
PT the material.
XX Example 3; Fig 24; 77pp; English.
PS The invention relates to a method of obtaining cell-wall material of a
XX Gram-positive bacterium with improved capacity for binding to a
CC proteinaceous substance (PS), or binding PS to cell-wall material of the
CC bacterium, by treating the cell-wall material with a solution capable of
CC removing a cell-wall component such as a protein, (lipo)teichoic acid or
CC carbohydrate from the material. The method especially uses the binding of
CC the PS with an anchor protein from the cell wall of the organism. The
CC anchor protein is preferably the Lactococcus lactis AcmaA cell wall
CC binding domain or its homolog or functional derivative. The method is
CC useful for improving binding of proteinaceous substance to cell wall
CC material of gram-positive bacterium, e.g. for the preparation of a
CC pharmaceutical composition comprising a vaccine useful for mucosal
CC immunization and for preparing a biocatalyst. The preparation is also
CC useful for generating bioadsorbents or biofilters for environmental
CC purposes, microbiocatalysts and diagnostic tools. Also for vaccination
CC Streptococcus pneumoniae. In an example of the invention, acid
CC pretreatment of Gram-positive bacteria to enhance binding of AcmaA protein
CC anchor fusion is investigated by the construction of a reporter plasmid
CC encoding a fusion gene of the L. lactis AcmaD protein and the
CC Streptococcus pneumoniae protease maturation protein PpmA. This sequence
CC represents a homologue of the Lactococcus lactis AcmaA anchor protein. The
CC region can be used in the generation of the fusion protein of the
CC invention
XX
SQ Sequence 44 AA;
Query Match 17.1%; Score 47; DB 6; Length 44;
Best Local Similarity 32.6%; Pred. No. 2.5e+02;
Matches 15; Conservative 6; Mismatches 21; Indels 4; Gaps 2;
QY 5 YTTDGNVTSSDGLWNNNTQFLFLEHSLLTANTTKGIYAGVCIS 50
DB 1 YVVKQDITLSGIASN-WGTNWQELARQNSLSPNM---IYAGQVIS 42
RESULT 14
AD70020
ID ADF70020 standard; protein; 44 AA.
XX

AC ADF70020;
XX 12-FEB-2004 (first entry)
XX AcmaA-type homologue amino acid sequence.
DE delivery; targeting system; AcmaA-type anchor protein; solid tumour;
KW health; medical; agricultural; cosmetic; controlled release.
XX Unidentified.
XX WO2003084508-A1.
FN 16-OCT-2003.
XX 04-APR-2003; 2003WO-NL000256.
PF 04-APR-2002; 2002EP-00076316.
XX 04-APR-2002; 2002US-0369927P.
PR 05-APR-2002; 2002US-0370485P.
PR 20-DEC-2002; 2002EP-00080481.
XX (NANO-) APPLIED NANOSYSTEMS BV.
PA Friesen RHE, Leenhouts CJ, Hektor HJ, Van Esch JH, Heeres A;
XX Robillard GT;
PI WPI; 2003-877005/81.
DR Vehicle for targeted delivery of therapeutic or diagnostic agents,
XX PT includes protein anchor and system for inducing availability of the
PT agent.
XX Example 3; Page 193; 303pp; English.
PS The present invention describes a vehicle (A) for delivering a substance
XX (I) to a predetermined site, which comprises (I); a system for inducing
CC availability of at least one compartment of (A) towards the exterior;
CC and, as targeting system for directing (A) to the site, an AcmaA-type
CC anchor protein (II). (A) are used for delivery of diagnostic and
CC therapeutic agents to predetermined sites in the body, particularly
CC joints or solid tumours but can be used more generally for health,
CC medical, agricultural and cosmetic applications. (A) significantly
CC increases the half-life of peptides in the circulation and, by providing
CC controlled release, ensures relatively high bioavailability, allowing
CC therapeutic use of agents that would otherwise be too toxic for systemic
CC administration. The native AcmaA peptide targets Gram-positive bacteria
CC but its homologues can be engineered to have different selectivity. The
CC present sequence is used in the exemplification of the present invention.
XX
SQ Sequence 44 AA;
Query Match 17.1%; Score 47; DB 7; Length 44;
Best Local Similarity 32.6%; Pred. No. 2.5e+02;
Matches 15; Conservative 6; Mismatches 21; Indels 4; Gaps 2;
QY 5 YTTDGNVTSSDGLWNNNTQFLFLEHSLLTANTTKGIYAGVCIS 50
DB 1 YVVKQDITLSGIASN-WGTNWQELARQNSLSPNM---IYAGQVIS 42
RESULT 15
AAM92081
ID AAM92081 standard; protein; 49 AA.
XX AAM92081;
AC 06-NOV-2001 (first entry)
XX Human digestive system antigen SEQ ID NO: 1430.
DE Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;

KW digestive system disorder; Meckel's diverticulum.

XX	Homo sapiens.	PR	25-SEP-2000; 2000US-0234998P.
OS	WO20015314-A2.	PR	26-SEP-2000; 2000US-0235484P.
XX	02-AUG-2001.	PR	27-SEP-2000; 2000US-0235834P.
XX	17-JAN-2001; 2001WO-US001324.	PR	27-SEP-2000; 2000US-0235836P.
XX	31-JAN-2000; 2000US-0179065P.	PR	29-SEP-2000; 2000US-0236327P.
XX	04-FEB-2000; 2000US-0180628P.	PR	29-SEP-2000; 2000US-0236367P.
XX	24-FEB-2000; 2000US-0184664P.	PR	29-SEP-2000; 2000US-0236368P.
XX	02-MAR-2000; 2000US-0186350P.	PR	29-SEP-2000; 2000US-0236369P.
XX	16-MAR-2000; 2000US-0189874P.	PR	29-SEP-2000; 2000US-0236370P.
XX	17-MAR-2000; 2000US-0190078P.	PR	02-OCT-2000; 2000US-0236802P.
XX	18-APR-2000; 2000US-0198123P.	PR	02-OCT-2000; 2000US-0237037P.
XX	19-MAY-2000; 2000US-0205515P.	PR	02-OCT-2000; 2000US-0237038P.
XX	07-JUN-2000; 2000US-0209467P.	PR	02-OCT-2000; 2000US-0237039P.
XX	28-JUN-2000; 2000US-0214886P.	PR	13-OCT-2000; 2000US-0237040P.
XX	30-JUN-2000; 2000US-0215135P.	PR	13-OCT-2000; 2000US-0239935P.
XX	07-JUL-2000; 2000US-0216647P.	PR	13-OCT-2000; 2000US-0239937P.
XX	07-JUL-2000; 2000US-0216880P.	PR	20-OCT-2000; 2000US-0240960P.
XX	11-JUL-2000; 2000US-0217487P.	PR	20-OCT-2000; 2000US-0241221P.
XX	11-JUL-2000; 2000US-0217496P.	PR	20-OCT-2000; 2000US-0241785P.
XX	14-JUL-2000; 2000US-0218290P.	PR	20-OCT-2000; 2000US-0241786P.
XX	26-JUL-2000; 2000US-0220963P.	PR	20-OCT-2000; 2000US-0241787P.
XX	26-JUL-2000; 2000US-0220964P.	PR	20-OCT-2000; 2000US-0241808P.
XX	14-AUG-2000; 2000US-0224518P.	PR	20-OCT-2000; 2000US-0241809P.
XX	14-AUG-2000; 2000US-0224519P.	PR	01-NOV-2000; 2000US-0241826P.
XX	14-AUG-2000; 2000US-0225213P.	PR	08-NOV-2000; 2000US-0244617P.
XX	14-AUG-2000; 2000US-0225214P.	PR	08-NOV-2000; 2000US-0246474P.
XX	14-AUG-2000; 2000US-0225266P.	PR	08-NOV-2000; 2000US-0246475P.
XX	14-AUG-2000; 2000US-0225267P.	PR	08-NOV-2000; 2000US-0246476P.
XX	14-AUG-2000; 2000US-0225268P.	PR	08-NOV-2000; 2000US-0246477P.
XX	14-AUG-2000; 2000US-0225270P.	PR	08-NOV-2000; 2000US-0246478P.
XX	14-AUG-2000; 2000US-0225447P.	PR	08-NOV-2000; 2000US-0246523P.
XX	14-AUG-2000; 2000US-0225757P.	PR	08-NOV-2000; 2000US-0246524P.
XX	14-AUG-2000; 2000US-0225758P.	PR	08-NOV-2000; 2000US-0246525P.
XX	14-AUG-2000; 2000US-0225759P.	PR	08-NOV-2000; 2000US-0246526P.
XX	18-AUG-2000; 2000US-0226279P.	PR	08-NOV-2000; 2000US-0246527P.
XX	22-AUG-2000; 2000US-0226681P.	PR	08-NOV-2000; 2000US-0246532P.
XX	22-AUG-2000; 2000US-0226868P.	PR	08-NOV-2000; 2000US-0246609P.
XX	22-AUG-2000; 2000US-0227182P.	PR	08-NOV-2000; 2000US-0246610P.
XX	23-AUG-2000; 2000US-0227003P.	PR	08-NOV-2000; 2000US-0246611P.
XX	30-AUG-2000; 2000US-0228924P.	PR	08-NOV-2000; 2000US-0246613P.
XX	01-SEP-2000; 2000US-0229287P.	PR	08-NOV-2000; 2000US-0249207P.
XX	01-SEP-2000; 2000US-0229343P.	PR	17-NOV-2000; 2000US-0249208P.
XX	01-SEP-2000; 2000US-0229345P.	PR	17-NOV-2000; 2000US-0249209P.
XX	05-SEP-2000; 2000US-0229503P.	PR	17-NOV-2000; 2000US-0249210P.
XX	05-SEP-2000; 2000US-0229513P.	PR	17-NOV-2000; 2000US-0249211P.
XX	06-SEP-2000; 2000US-0230437P.	PR	17-NOV-2000; 2000US-0249212P.
XX	06-SEP-2000; 2000US-0230438P.	PR	17-NOV-2000; 2000US-0249218P.
XX	08-SEP-2000; 2000US-0231242P.	PR	17-NOV-2000; 2000US-0249244P.
XX	08-SEP-2000; 2000US-0231243P.	PR	17-NOV-2000; 2000US-0249245P.
XX	08-SEP-2000; 2000US-0231244P.	PR	17-NOV-2000; 2000US-0249264P.
XX	08-SEP-2000; 2000US-0231413P.	PR	17-NOV-2000; 2000US-0249265P.
XX	08-SEP-2000; 2000US-0231414P.	PR	17-NOV-2000; 2000US-0249297P.
XX	08-SEP-2000; 2000US-0232080P.	PR	17-NOV-2000; 2000US-0249300P.
XX	08-SEP-2000; 2000US-0232081P.	PR	01-DEC-2000; 2000US-0250160P.
XX	12-SEP-2000; 2000US-0231968P.	PR	01-DEC-2000; 2000US-0250391P.
XX	14-SEP-2000; 2000US-0232397P.	PR	05-DEC-2000; 2000US-0251030P.
XX	14-SEP-2000; 2000US-0232398P.	PR	05-DEC-2000; 2000US-0251988P.
XX	14-SEP-2000; 2000US-0232399P.	PR	05-DEC-2000; 2000US-0256719P.
XX	14-SEP-2000; 2000US-0232400P.	PR	06-DEC-2000; 2000US-0251479P.
XX	14-SEP-2000; 2000US-0232401P.	PR	08-DEC-2000; 2000US-0251856P.
XX	14-SEP-2000; 2000US-0233063P.	PR	08-DEC-2000; 2000US-0251858P.
XX	14-SEP-2000; 2000US-0233064P.	PR	08-DEC-2000; 2000US-0251869P.
XX	14-SEP-2000; 2000US-0233065P.	PR	08-DEC-2000; 2000US-0251989P.
XX	21-SEP-2000; 2000US-0234223P.	PR	08-DEC-2000; 2000US-0251990P.
XX	21-SEP-2000; 2000US-0234274P.	PR	11-DEC-2000; 2000US-0254097P.
XX	25-SEP-2000; 2000US-0234997P.	PR	05-JAN-2001; 2001US-02559678P.


```
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-502630/55.
DR N-PSDB; AAK87854.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
PS Claim 11; SEQ ID NO 1430; 986pp; English.
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a digestive system antigen of
CC the invention
XX
SQ Sequence 49 AA;
Query Match 16.7%; Score 46; DB 4; Length 49;
Best Local Similarity 33.3%; Pred. No. 3.8e+02;
Matches 9; Conservative 3; Mismatches 15; Indels 0; Gaps 0;
QY 2 LVSYYTGDGNDVTVESDGLWNNNTQL 28
Db 23 IMKYSTQGDQKEXEGKASLWNLHTWL 49
Search completed: June 29, 2005, 09:41:08
Job time : 36.5521 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 20.8409 Seconds
(without alignments)

1277.688 Million cell updates/sec

Title: US-10-718-321-7_COPY_249_300

Perfect score: 275

Sequence: 1 PLYSYTDCNDVTSSDGL.....SLLTANTTKGIYAGVCISVL 52

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	275	100.0	359	2	O43656	O43656 homo sapien
2	275	100.0	364	2	Q96D42	Q96D42 homo sapien
3	228	82.9	451	2	Q95144	Q95144 cercopithec
4	228	82.9	460	2	O18984	O18984 cercopithec
5	228	82.9	469	2	Q7J348	Q7J348 cercopithec
6	228	82.9	473	2	Q7J347	Q7J347 cercopithec
7	228	82.9	474	2	O46597	O46597 cercopithec
8	228	82.9	478	2	O46598	O46598 cercopithec
9	79	28.7	305	2	Q8VIM2	Q8VIM2 mus musculus
10	78	28.4	307	2	O54947	O54947 rattus norv
11	76.5	27.8	282	2	O8VIM1	O8VIM1 mus musculus
12	76.5	27.8	282	2	Q7TPU2	Q7TPU2 mus musculus
13	68	24.7	549	2	Q53969	Q53969 shigella dy
14	67.5	24.5	460	2	Q7URM5	Q7URM5 rhodopirell
15	65	23.6	554	2	Q842D0	Q842D0 escherichia
16	65	23.6	554	2	O9KJA0	O9KJA0 escherichia
17	63	22.9	1687	2	Q7QIM9	Q7QIM9 anopheles g
18	62.5	22.7	639	1	YHR7_YEAST	YHR7_YEAST
19	62.5	22.7	931	2	Q6WP59	Q6WP59 dengue viru
20	62.5	22.7	931	2	Q6WP60	Q6WP60 dengue viru
21	62.5	22.7	931	2	Q6WP63	Q6WP63 dengue viru
22	62.5	22.7	931	2	Q6WP68	Q6WP68 dengue viru
23	62.5	22.7	931	2	Q6WP69	Q6WP69 dengue viru
24	62.5	22.7	931	2	Q6WP70	Q6WP70 dengue viru
25	62.5	22.7	1269	2	Q97206	Q97206 sulfolobus
26	62.5	22.7	3392	2	Q8JQD9	Q8JQD9 dengue viru
27	62.5	22.7	3392	2	Q91NH1	Q91NH1 dengue viru
28	62.5	22.7	3392	2	Q6WP58	Q6WP58 dengue viru
29	62.5	22.7	3392	2	Q80KC7	Q80KC7 dengue viru
30	62	22.5	329	1	SRAE_CABEEL	Q09208 caenorhabdi
31	62	22.5	411	2	Q23351	Q23351 caenorhabdi

32	62	22.5	472	1	VL2_HPV34	P36758 human papil
33	62	22.5	842	2	Q9U0J9	Q9U0J9 plasmodium
34	61	22.2	133	2	Q6GZ26	Q6GZ26 fremyella d
35	61	22.2	1255	2	O36287	O36287 venezuelan
36	60.5	22.0	103	1	CYC_THUAA	P81459 thunnus ala
37	60.5	22.0	126	2	Q9P819	Q9P819 erwinia amy
38	60.5	22.0	474	2	Q9H751	Q9H751 homo sapien
39	60.5	22.0	474	2	Q9HC44	Q9HC44 homo sapien
40	60.5	22.0	3056	2	Q7USQ0	Q7USQ0 rhodopirell
41	60	21.8	282	2	Q7RJE4	Q7RJE4 plasmodium
42	60	21.8	287	2	O52953	O52953 bacillus su
43	60	21.8	338	2	Q9SQ06	Q9SQ06 arabidopsis
44	60	21.8	602	1	NRGI_CHICK	Q05199 gallus gall
45	60	21.8	1850	2	Q75CI2	Q75CI2 ashbya goss

ALIGNMENTS

RESULT 1		PRELIMINARY;		PRT;		359 AA.	
O43656		AC O43656;		DT 01-JUN-1998		(TrEMBLrel. 06, Created)	
		DT 01-JUN-1998		(TrEMBLrel. 06, Last sequence update)			
		DT 01-MAR-2004		(TrEMBLrel. 26, Last annotation update)			
		DE Hepatitis A virus cellular receptor 1.					
		GN Name=HAVcr-1;					
		OS Homo sapiens (Human).					
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
		OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
		OX NCBI_TaxID=9606;					
		RN [1]					
		RP SEQUENCE FROM N.A.					
		RC TISSUE=Liver;					
		RX MEDLINE=98325180; PubMed=9658108;					
		RA Feigelsstock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;					
		RT "The human homolog of HAVcr-1 codes for a hepatitis A virus cellular					
		RT receptor.";					
		RL J. Virol. 72:6621-6628(1998).					
		DR EMBL; AF043724; AAC39862.1; -					
		DR GO; GO:0004872; Fireceptor activity; IEA.					
		DR InterPro; IPR003599; IG.					
		DR InterPro; IPR007110; IG-like.					
		DR InterPro; IPR003006; IG_MHC.					
		DR SMART; SM00409; IG; 1.					
		DR PROSITE; PS00835; IG LIKE; 1.					
		DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.					
		KW Receptor.					
		SQ SEQUENCE 359 AA; 38704 MW; C207FEC562DC62CA CRC64;					
		Query Match 100.0%; Score 275; DB 2; Length 359;					
		Best Local Similarity 100.0%; Pred. No. 3e-25; Indels 0; Gaps 0;					
		Matches 52; Conservative 0; Mismatches 0;					
QY		1 PLYSYTDCNDVTSSDGLWNNNQQLFLEHSLLTANTTKGIYAGVCISVL 52					
Db		249 PLYSYTDCNDVTSSDGLWNNNQQLFLEHSLLTANTTKGIYAGVCISVL 300					
RESULT 2		PRELIMINARY;		PRT;		364 AA.	
Q96D42		AC Q96D42;		DT 01-DEC-2001		(TrEMBLrel. 19, Created)	
		DT 01-DEC-2001		(TrEMBLrel. 19, Last sequence update)			
		DT 25-OCT-2004		(TrEMBLrel. 28, Last annotation update)			
		DE HAVCR1 protein.					
		GN Name=HAVCR1;					
		OS Homo sapiens (Human).					
		OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
		OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
		OX NCBI_TaxID=9606;					


```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea;
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4."
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043446; AAC39771.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 469 AA; 50561 MW; 27E386F47948F528 CRC64;

Query Match 82.9%; Score 228; DB 2; Length 469;
Best Local Similarity 84.6%; Pred. No. 2.4e-19;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTNDGNTVTSSDGLMNNQTLFLEHSLTANTTKGIYAGVCISVL 52
Db 347 PLYSYTDDGSDTVTSSDGLMNNQTLFLEHSLTANTTKGIYAGVCISVL 398

RESULT 6
Q7JJ47
ID Q7JJ47 PRELIMINARY; PRT; 473 AA.
AC Q7JJ47
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hepatitis A virus cellular receptor 1 short form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea;
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4."
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043448; AAC39773.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 473 AA; 50973 MW; CD15EF5EE79C8013 CRC64;

Query Match 82.9%; Score 228; DB 2; Length 473;
Best Local Similarity 84.6%; Pred. No. 2.4e-19;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTNDGNTVTSSDGLMNNQTLFLEHSLTANTTKGIYAGVCISVL 52
Db 351 PLYSYTDDGSDTVTSSDGLMNNQTLFLEHSLTANTTKGIYAGVCISVL 402

us-10-718-321-7_copy_249_300.rup

```

```

RESULT 7
O46597
ID O46597 PRELIMINARY; PRT; 474 AA.
AC O46597
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hepatitis A virus cellular receptor 1 long form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea;
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4."
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043447; AAC39772.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 474 AA; 51088 MW; 472DF987CA220524 CRC64;

Query Match 82.9%; Score 228; DB 2; Length 474;
Best Local Similarity 84.6%; Pred. No. 2.4e-19;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTNDGNTVTSSDGLMNNQTLFLEHSLTANTTKGIYAGVCISVL 52
Db 352 PLYSYTDDGSDTVTSSDGLMNNQTLFLEHSLTANTTKGIYAGVCISVL 403

RESULT 8
O46598
ID O46598 PRELIMINARY; PRT; 478 AA.
AC O46598
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hepatitis A virus cellular receptor 1 long form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecoidea;
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4."
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043449; AAC39774.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG_1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.

```

```
KW Receptor.
SQ SEQUENCE 478 AA; 51501 NW; 3A919655C752FF60 CRC64;

Query Match 82.9%; Score 228; DB 2; Length 478;
Best Local Similarity 84.6%; Pred. No. 2.4e-19;
Matches 44; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 PLYSYTTDGNVTVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 52
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 356 PLYSYTTDGSDFVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 407
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 9
QBVM2
ID QBVM2 PRELIMINARY; PRT; 305 AA.
AC QBVM2;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE TIM1.
CN Name=Havcr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BA1B/c; TISSUE=Spleen;
RX MEDLINE=21582130; PubMed=11725301;
RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Kuchroo V.K.,
RA Barsh G.S., Freeman G.J., Umetsu D.T., DeKruyff R.H.;
RT "Identification of Tapr (an airway hyperreactivity regulatory locus)
RT and the linked Tim gene family.";
RL Nat. Immunol. 2:1109-1116(2001).
DR EMBL; AF399829; AAL35774.1; -.
DR MGD; MGI:2159680; Havcr1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR SEQUENCE 305 AA; 33391 NW; 8FAEA38627FE85FB CRC64;

Query Match 28.7%; Score 79; DB 2; Length 305;
Best Local Similarity 42.0%; Pred. No. 0.27;
Matches 21; Conservative 6; Mismatches 19; Indels 4; Gaps 3;

OY 4 SYT-TDGNVTVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 52
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 201 SHTPTDNGTVTSSGD-TWSNHTEA--IPPGKPKQKPTKGFYVGICIAAL 247
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 10
O54947
ID O54947 PRELIMINARY; PRT; 307 AA.
AC O54947;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Kidney injury molecule-1 (KIM-1 protein).
CN Name=KIM-1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Kidney;
RX MEDLINE=98129827; PubMed=9461608; DOI=10.1074/jbc.273.7.4135;
RA Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,
RA Cate R.L., Sanicola M.;
RT "Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion
RT molecule containing a novel immunoglobulin domain, is up-regulated in
RT renal cells after injury.";
```

```
J. Biol. Chem. 273:4135-4142(1998).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Heiton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF035963; AAC53546.1; -.
DR EMBL; BC061820; AAH61820.1; -.
DR InterPro; IPR003599; IG.
DR SMART; SM007110; IG-like.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR SEQUENCE 307 AA; 33963 MW; 736D1DD1F1549760 CRC64;

Query Match 28.4%; Score 78; DB 2; Length 307;
Best Local Similarity 42.0%; Pred. No. 0.36;
Matches 21; Conservative 7; Mismatches 18; Indels 4; Gaps 3;

OY 4 SYT-TDGNVTVTSSDGLWNNNTQFLFHSLLTANTTKGIYAGVCISVL 52
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
DB 199 SYTADWNGTWT-SSEAWNNHTVRIPLRKP--QRNPTKGFYVGMSVAAL 245
|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

RESULT 11
QBVM1
ID QBVM1 PRELIMINARY; PRT; 282 AA.
AC QBVM1;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE TIM1.
CN Name=Havcr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2; TISSUE=Spleen;
RX MEDLINE=21582130; PubMed=11725301;
RA McIntire J.J., Umetsu S.E., Akbari O., Potter M., Kuchroo V.K.,
RA Barsh G.S., Freeman G.J., Umetsu D.T., DeKruyff R.H.;
RT "Identification of Tapr (an airway hyperreactivity regulatory locus)
RT and the linked Tim gene family.";
RL Nat. Immunol. 2:1109-1116(2001).
DR EMBL; AF399830; AAL35775.1; -.
DR MGD; MGI:2159680; Havcr1.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
```

```
DR SMART; SM00409; IG; 1.
DR PROSITE; PS0835; IG LIKE; 1.
SQ SEQUENCE 282 AA; 30968 MW; 7D30ER0698FOAC5F CRC64;

Query Match
Best Local Similarity 27.8%; Score 76.5; DB 2; Length 282;
Matches 18; Conservative 8; Mismatches 21; Indels 3; Gaps 2;

QY 3 YSYTTDGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 178 WTHKPDWNGTIVSSGD-TWSNHTA--IPPGKPKQNKPTKGFVVICIAAL 224

RESULT 12
Q7TPU2 PRELIMINARY; PRT; 282 AA.
AC Q7TPU2;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Blasctocyst;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Roschlyuk S., Carrincci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Blasctocyst;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053400; AAH53400.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS0835; IG LIKE; 1.
KW Hypothetical protein.
SQ SEQUENCE 282 AA; 30938 MW; 6D31E2778480AC5F CRC64;

Query Match
Best Local Similarity 27.8%; Score 76.5; DB 2; Length 282;
Matches 18; Conservative 8; Mismatches 21; Indels 3; Gaps 2;

QY 3 YSYTTDGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISVL 52
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 178 WTHKPDWNGTIVSSGD-TWSNHTA--IPPGKPKQNKPTKGFVVICIAAL 224

RESULT 13
Q53969 PRELIMINARY; PRT; 549 AA.
AC Q53969;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Flagellin.
GN Name=flc-SD;
OS Shigella dysenteriae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=622;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sh16;
RA Mahmoud M.A., Tominaga A., Abu Amar A.M., Enomoto M.;
RL Submitted (DEC-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL; D26166; BAA05153.1; -.
DR HSP; P06179; IUCU.
DR GO; GO:0009420; C:flagellar filament (sensu Bacteria); IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0001539; P:ciliary or flagellar motility; IEA.
DR InterPro; IPR001029; Flagellin_C.
DR InterPro; IPR001492; Flagellin_N.
DR Pfam; PF00700; Flagellin_C; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD000316; Flagellin_C; 1.
KW Flagellum.
SQ SEQUENCE 549 AA; 57591 MW; 8C8A56DC661CB732 CRC64;

Query Match
Best Local Similarity 24.7%; Score 68; DB 2; Length 549;
Matches 21; Conservative 7; Mismatches 18; Indels 4; Gaps 2;

QY 4 SYTTDGDNDVTVESSDGLWNNQTLFLEHSLLTANTTKGIYAGVCISV 51
   ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 322 SKTNAGTQAKLSLMANNANAKVTITDKGTFNTANTIK--FDGVDISV 369

RESULT 14
Q7UKM5 PRELIMINARY; PRT; 460 AA.
AC Q7UKM5;
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Similar to Cap5J protein-putative transmembrane protein.
GN Ordered locus names=RL10055;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294150; CAD76607.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR007016; Wzy_C.
DR Pfam; PF04932; Wzy_C; 1.
KW Complete proteome; Transmembrane.
SQ SEQUENCE 460 AA; 48823 MW; CB4F7122CE5A7E1B CRC64;

Query Match
Best Local Similarity 24.5%; Score 67.5; DB 2; Length 460;
Matches 17; Conservative 5; Mismatches 10; Indels 7; Gaps 2;
```

Qy 21 WNNQTLFLE-----HSLLT-ANTTKGIYAGVCISVL 52
Db 346 WGLNSAPLLEDFSYHTHLLHLAIFSAGIFAGICVSVL 384

RESULT 15

Q842D0
ID Q842D0 PRELIMINARY; PRT; 554 AA.
AC Q842D0;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE FLIC.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K 12a;
RX MEDLINE=22586428; PubMed=12700273;
RX DOI=10.1128/JB.185.9.2936-2943.2003;
RA Wang L., Rothermund D., Curd H., Reeves P.R.;
RT "Species-wide variation in the Escherichia coli flagellin (H-antigen)
gene";
RL J. Bacteriol. 185:2936-2943 (2003).
DR EMBL; AY250001; AAP13309.1; -.
DR HSSP; P06179; 1UCU.
DR GO; GO:0009420; C:flagellar filament (sensu Bacteria); IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0001539; P:ciliary or flagellar motility; IEA.
DR InterPro; IPR001029; Flagellin C.
DR InterPro; IPR001492; Flagellin_N.
DR Pfam; PF00700; Flagellin_C; 1.
DR Pfam; PF00669; Flagellin_N; 1.
DR PRINTS; PR00207; FLAGELLIN.
DR ProDom; PD00316; Flagellin_C; 1.
SQ SEQUENCE 554 AA; 57243 MW; 3CF9E93C1612BF91 CRC64;

Query Match 23.6%; Score 65; DB 2; Length 554;
Best Local Similarity 42.0%; Pred.No. 28;
Matches 21; Conservative 6; Mismatches 19; Indels 4; Gaps 2;

Qy 4 SVTTDGNLTVTSSDGLWN--NNQTLFLEHSLLTANTTKGIYAGVCISV 51
Db 327 SKTNAGTTQAKLSLDMANNANAKVITTDKGTFTANTTK--FDGVDISV 374

Search completed: June 29, 2005, 09:01:46
Job time : 21.8409 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 57.6071-Seconds
(without alignments)
543.815 Million cell updates/sec

Title: US-10-718-321-6

Perfect score: 432

Sequence: 1 MFLPRQNHVPATSPSPQP.....NNQTQLFLESLTANTTKG 81

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_16Dec04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	432	100.0	81	6	AAO26678 Human KIM
2	432	100.0	298	8	ADQ76693 Human KIM
3	432	100.0	334	2	AAW38336 Human kid
4	432	100.0	334	6	AAO26679 KIM-1 rel
5	432	100.0	334	8	ADQ76696 Human kid
6	432	100.0	339	7	ADQ76618 Plasmid p
7	432	100.0	339	7	ADQ766592 Human NOV
8	432	100.0	359	6	ABR58582 Human can
9	432	100.0	359	6	ABP70439 Amino aci
10	432	100.0	359	6	ABP70438 Amino aci
11	432	100.0	359	6	ABR48174 Human bla
12	432	100.0	359	6	AAO26680 KIM-1 rel
13	432	100.0	359	7	ADQ76594 Human NOV
14	432	100.0	359	7	ADN38984 Cancer/an
15	432	100.0	359	8	ADQ76690 Human kid
16	432	100.0	364	6	ABP70442 Amino aci
17	432	100.0	365	6	ABP70440 Amino aci
18	432	100.0	518	8	ADQ76691 Human KIM
19	427	98.8	364	6	ABP70443 Amino aci
20	423	97.9	359	6	ABP70441 Amino aci
21	396	91.7	263	7	ADQ76629 Human NOV
22	396	91.7	263	7	ADQ76622 Plasmid p
23	396	91.7	263	7	ADQ76596 Human NOV
24	348	80.6	451	2	AAR92803 Hepatitis
25	101	23.4	18	6	AAO26686 Human KIM

26	99	22.9	18	6	AAO26687	Aao26687 Human KIM
27	99	22.9	18	6	AAO26673	Aao26673 Monoclonal
28	98	22.7	18	6	AAO26681	Aao26681 Human KIM
29	97	22.5	18	6	AAO26682	Aao26682 Human KIM
30	96	22.2	18	6	AAO26683	Aao26683 Human KIM
31	95	22.0	18	6	AAO26685	Aao26685 Human KIM
32	92	21.3	18	6	AAO26684	Aao26684 Human KIM
33	90	20.8	18	6	AAO26688	Aao26688 Human KIM
34	89	20.6	307	2	AAW38334	Aaw38334 Rat kidne
35	88	20.4	153	3	AAB42784	Aab42784 Human ORF
36	88	20.4	153	5	ABP69679	Abp69679 Human pol
37	79	18.3	681	5	ABP93650	Abp93650 Herbicida
38	76.5	17.7	282	6	ABP70431	Abp70431 Amino aci
39	76	17.6	447	3	AAB42121	Aab42121 Human ORF
40	76	17.6	647	3	AAV53000	AAv53000 Human sec
41	76	17.6	893	7	ADK40920	Adk40920 Novel hum
42	76	17.6	893	8	ADR15641	Adr15641 Kinase 42
43	75	17.4	8360	6	AAE35499	Aae35499 Streptomy
44	74.5	17.2	539	6	ABP78670	Abp78670 N. gonorr
45	72.5	16.8	219	3	AAG19395	Aag19395 Arabidops

ALIGNMENTS

RESULT 1
AAO26678
ID AAO26678 standard; protein; 81 AA.
XX
AC AAO26678;
XX
DT 20-MAR-2003 (first entry)
XX
DE Human KIM-1 mucin domain protein, SEQ ID No 6.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX
OS Homo sapiens.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
(BIOJ) BIOGEN INC.
(GEHO) GEN HOSPITAL CORP.
XX
PI Bailly V, Bonventre J;
XX
DR WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
e.g. renal cancer.
XX
PS Disclosure; Fig 1A; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a human KIM-1 mucin domain protein
XX of the invention

SQ Sequence 81 AA;
 Query Match 100.0%; Score 432; DB 6; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.7e-36;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTSSDGLW 60
 DB 1 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTSSDGLW 60

QY 61 NNNQTQLFLEHSLLTANTTKG 81
 DB 61 NNNQTQLFLEHSLLTANTTKG 81

RESULT 2
 ADQ76693
 ID ADQ76693 standard; protein; 298 AA.
 XX
 AC ADQ76693;
 XX
 DT 07-OCT-2004 (first entry)
 XX Human KIM-1 extracellular domain-6xHis fusion.
 DE
 XX Human; KIM-1; kidney injury module-1; immunosuppressive; protozoasides;
 KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
 KW antiallergic; antiinflammatory; vasotropic; nephrotropic.
 XX
 OS Homo sapiens.
 OS Synthetic.

Key Location/Qualifiers
 Peptide 1..20
 /label= signal_peptide

WO2004060041-A2.
 22-JUL-2004.
 29-DEC-2003; 2003WO-US041294.
 30-DEC-2002; 2002US-0436934P.
 (BIOG-) BIOGEN IDEC MA INC.
 Rennett PD;
 WPI; 2004-534277/51.

Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell activation, or inhibiting production of a subset of antibodies against one or more antigens.

Example 3; Page 33; 68pp; English.

The present sequence is that of a polypeptide comprising the extracellular domain (residues 1-290) of human kidney injury molecule-1 (KIM-1) ADQ76690 fused to a C-terminal 6xHis tag peptide. A plasmid comprising DNA encoding this sequence under control of a CMV promoter was used for transient constitutive expression of the polypeptide in mammalian cells in an example from the invention. The invention provides methods for therapeutically modulating immune function in autoimmune diseases and other disorders of the mammalian immune system. A method is claimed for inhibiting signalling between a T cell and a second cell, e.g. an antigen-presenting cell, in a mammal. The method involves identifying a mammal with an immune disease or disorder, or one preparing to receive a tissue graft, and administering a KIM-1 antagonist, especially a polypeptide comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a soluble polypeptide, which can

CC include a KIM-1 mucin domain in addition to the KIM Ig domain, and may include a heterologous moiety such as an Fc moiety. The antagonist may be conjugated to a polymer such as polyethylene glycol. The KIM-1 antagonist is used in methods of: inhibiting activation of a B cell in a mammal; inhibiting disease relapse in an autoimmune disease; inhibiting epitope spreading in an autoimmune disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis, autoimmune haemolytic anaemia, Chagas disease, Grave's disease, idiopathic thrombocytopenia purpura, Wegener's granulomatosis, polyarteritis nodosa, rapidly progressive crescentic glomerulonephritis, graft-versus-host disease, or systemic lupus nephritis) or inflammatory disease or disorder (e.g. inflammatory bowel diseases such as Crohn's disease, ulcerative colitis, and ileitis); and inhibiting secretion of IFN-gamma by lymphocytes.

SQ Sequence 298 AA;
 Query Match 100.0%; Score 432; DB 8; Length 298;
 Best Local Similarity 100.0%; Pred. No. 8.1e-36;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTSSDGLW 60
 DB 210 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTSSDGLW 269

QY 61 NNNQTQLFLEHSLLTANTTKG 81
 DB 270 NNNQTQLFLEHSLLTANTTKG 290

RESULT 3
 AA38336
 ID AA38336 standard; protein; 334 AA.
 XX
 AC AA38336;
 XX
 DT 21-MAY-1998 (first entry)
 XX Human kidney injury related molecule (KIM).
 DE
 XX Kidney injury related molecule; KIM; human; renal disease; injury;
 KW nephritis; tissue regeneration; therapy; monoclonal antibody.
 XX
 OS Homo sapiens.
 XX
 PN WO9744460-A1.
 XX
 PD 27-NOV-1997.
 XX
 PF 23-MAY-1997; 97WO-US009303.
 XX
 PR 24-MAY-1996; 96US-0018228P.
 PR 23-AUG-1996; 96US-0023442P.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Sanicola-Nadel M, Bonventre JV, Hession CA, Ichimura T, Wei H;
 PI Cate RL;
 XX
 DR WPI; 1998-018514/02.
 XX
 PT DNA encoding kidney injury related molecule - which is upregulated in injured or regenerating tissue, useful to promote growth of new tissue and survival of damaged tissue.
 PT
 XX Claim 9; Page 46-47; 68pp; English.
 PS
 CC This protein, designated kidney injury related molecule (KIM), is up-regulated in injured or regenerating tissue. Its amino acid sequence was deduced from a clone (see AAT96035) obtained from a human embryonic liver library. A 572-amino acid (see AAW38335) and a 307-amino acid rat KIM (see AAW38334) are also claimed. Recombinant KIM polypeptides can be expressed in prokaryotic and eukaryotic host cells using a claimed process. Soluble variants fused to a toxin, imageable compound or

CC radionuclide, and IgG fusion proteins are also claimed. KIM, or an
CC agonist, can be used to treat renal disease and to promote the growth of new
CC tissue or the survival of damaged tissue, generally in conditions where
CC the binding of specific ligand to KIM stimulates cell growth, maintains
CC cellular differentiation or reduces apoptosis, e.g. in cases of renal
CC failure, nephritis, kidney transplants, toxic or hypoxic injury. A
CC monoclonal antibody specific for KIM can be used to treat renal disease,
CC e.g. where binding of KIM to ligand results in neoplasia, loss of
CC cellular function, susceptibility to apoptosis or promotion of
CC inflammation, deliver imaging agents to KIM expressing cells in vivo or
CC in vitro and measure KIM concentration by immunoassay.
CC Damage/regeneration of renal cells can be determined by measuring KIM,
CC particularly to diagnose or monitor the progress of disease or therapy.
CC KIM-expressing tumour cells can be inhibited by treatment with a fusion
CC protein comprising KIM ligand or mAb with a toxin or radionuclide, and
CC tumour cells that express KIM ligand can be inhibited with similarly
CC tagged KIM or anti-KIM ligand antibody
XX
SQ Sequence 334 AA;

Query Match 100.0%; Score 432; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 9.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPLPRQNHPEVATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNVTVESDGLW 60
DB 210 MPLPRQNHPEVATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNVTVESDGLW 269
QY 61 NNNQTQLFLEHSLTANTTKG 81
DB 270 NNNQTQLFLEHSLTANTTKG 290

RESULT 4
AAO26679
ID AAO26679 standard; protein; 334 AA.
XX
AC AAO26679;
XX
DT 20-MAR-2003 (first entry)
XX
DE KIM-1 related protein, SEQ ID No 7.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX
OS Homo sapiens.
XX
PN WO200298920-A1.
XX
PD 12-DEC-2002.
XX
PF 31-MAY-2002; 2002WO-US017402.
XX
PR 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
PA (BIOJ) BIOGEN INC.
PA (GEO) GEN HOSPITAL CORP.
XX
PI Baillly V, Bonventre J;
XX
DR WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Disclosure; Page 39-40; 42pp; English.
XX
CC The invention relates to a novel antibody, antibody derivative or antigen

CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a protein relating to the KIM-1
XX protein of the invention
SQ Sequence 334 AA;

Query Match 100.0%; Score 432; DB 6; Length 334;
Best Local Similarity 100.0%; Pred. No. 9.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MPLPRQNHPEVATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNVTVESDGLW 60
DB 210 MPLPRQNHPEVATSPSSQPAETHPTTLQGAIRREPTSSPLSYTTDGNVTVESDGLW 269
QY 61 NNNQTQLFLEHSLTANTTKG 81
DB 270 NNNQTQLFLEHSLTANTTKG 290

RESULT 5
ADQ76696
ID ADQ76696 standard; protein; 334 AA.
XX
AC ADQ76696;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human kidney injury molecule-1 (KIM-1).
XX
KW Human; KIM-1; kidney injury molecule-1; immunosuppressive; protozoacide;
KW muscular-gen.; neuroprotective; antinaemic; antithyroid; haemostatic;
KW antiallergic; antiinflammatory; vasotropic; nephrotropic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Modified-site /label= Signal_peptide
FT Modified-site /note= "N-glycosylated" 65..67
FT Modified-site /note= "N-glycosylated" 258..260
FT Modified-site /note= "N-glycosylated" 272..274
FT Modified-site /note= "N-glycosylated" 286..288
FT Modified-site /note= "N-glycosylated" 290..311
FT Domain /note= "Transmembrane domain"
FT Region 324..334
FT /note= "C-terminal region unique to this splice variant"
XX
PN WO2004060041-A2.
XX
PD 22-JUL-2004.
XX
PF 29-DEC-2003; 2003WO-US041294.
XX
PR 30-DEC-2002; 2002US-0436934P.
XX
PA (BIOG-) BIOGEN IDEC MA INC.
XX
PI Remmert PD;
XX
DR WPI; 2004-534277/51.
XX
PT Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
PT Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
PT activation, or inhibiting production of a subset of antibodies against
PT one or more antigens.

XX PS Disclosure; SEQ ID NO 1; 68pp; English.

XX CC The present sequence is that of human kidney injury molecule-1 (KIM-1), a

CC type I cell membrane glycoprotein and member of the T cell immunoglobulin

CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids or

CC 359 amino acids ADQ76690, depending on splice variation. It has been

CC discovered that treatment of a mammal with a KIM-1 antagonist alters the

CC interaction of T cells and other immune system cells, e.g. dendritic

CC cells, monocytes, macrophages and B cells, and thereby strongly

CC suppresses an IgG response to an antigen. Such treatment also eliminates

CC IgG1 production by memory B cells in response to subsequent challenge

CC with the antigen. Blockage of the binding of KIM-1 to its receptor

CC reduces secretion of interferon-gamma by immune cells engaged in an

CC antigen response in the mixed lymphocyte response assay. Based on these

CC discoveries, the invention provides methods for therapeutically

CC modulating immune function in autoimmune diseases and other disorders of

CC the mammalian immune system. A method is claimed for inhibiting

CC signalling between a T cell and a second cell, e.g. an antigen-presenting

CC cell, in a mammal. The method involves identifying a mammal with an

CC immune disease or disorder, or one preparing to receive a tissue graft,

CC and administering a KIM-1 antagonist, especially a polypeptide

CC comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a

CC KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding

CC fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a

CC soluble polypeptide, which can include a KIM-1 mucin domain in addition

CC to the KIM Ig domain, and may include a heterologous moiety such as an Fc

CC moiety. The antagonist may be conjugated to a polymer such as

CC polyethylene glycol. The KIM-1 antagonist is used in methods of:

CC inhibiting activation of a B cell in a mammal; inhibiting disease relapse

CC in an autoimmune disease; inhibiting epitope spreading in an autoimmune

CC disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis,

CC autoimmune haemolytic anaemia, Chagas disease, Grave's disease,

CC idiopathic thrombocytopenia purpura, Wegener's granulomatosis,

CC polyarteritis nodosa, rapidly progressive crescentic glomerulonephritis,

CC graft-versus-host disease, or systemic lupus nephritis) or inflammatory

CC disease or disorder (e.g. inflammatory bowel diseases such as Crohn's

CC disease, ulcerative colitis, and ileitis); and inhibiting secretion of

CC IFN-gamma by lymphocytes.

XX SQ Sequence 334 AA;

Query Match 100.0%; Score 432; DB 8; Length 334;

Best Local Similarity 100.0%; Pred. No. 9.3e-36;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDNDVTSSDGLW 60

Db 210 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDNDVTSSDGLW 269

Qy 61 NNNQTQLFLEHSLTANTTKG 81

Db 270 NNNQTQLFLEHSLTANTTKG 290

RESULT 6

ADE36618

ID ADE36618 standard; protein; 339 AA.

XX AC ADE36618;

XX DT 29-JAN-2004 (first entry)

XX DE Plasmid pCR2.1-CG57008-03-S843 15B protein insert SEQ ID NO:28.

XX KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;

XX KW renal cancer; inflammation; tissue typing.

XX OS Synthetic.

XX OS Homo sapiens.

XX FN WO2003080856-A2.

XX XX

PD 02-OCT-2003.

XX 19-MAR-2003; 2003WO-US008490.

XX 19-MAR-2002; 2002US-0365491P.

PR 13-SEP-2002; 2002US-0410618P.

XX (CURA-) CURAGEN CORP.

XX Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;

PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;

XX WPI; 2003-876927/81.

DR N-PSDB; ADE36617.

XX New polypeptide, useful for preparing a composition for treating or

PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer

PT or inflammation, or for tissue typing.

XX Example 1; SEQ ID NO 28; 239pp; English.

XX The present invention describes an isolated human NOVX polypeptide, where

CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at

CC least 95 % identical to it, or a sequence comprising one or more

CC conservative substitutions in the amino acid sequence. The NOVX

CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and

CC antiinflammatory activities, and can be used in gene therapy, and in

CC vaccines. The NOVX polypeptide is useful for preparing a composition for

CC treating or preventing a pathology associated with NOVX polypeptide e.g.

CC renal cancer or inflammation, or for tissue typing. The present sequence

CC represents a pCR2.1-CG57008-03-S843 15B protein insert, which is used in

CC an example from the present invention.

XX SQ Sequence 339 AA;

Query Match 100.0%; Score 432; DB 7; Length 339;

Best Local Similarity 100.0%; Pred. No. 9.5e-36;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDNDVTSSDGLW 60

Db 190 MPLPRQNHVPATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDNDVTSSDGLW 249

Qy 61 NNNQTQLFLEHSLTANTTKG 81

Db 250 NNNQTQLFLEHSLTANTTKG 270

RESULT 7

ADE36592

ID ADE36592 standard; protein; 339 AA.

XX AC ADE36592;

XX DT 29-JAN-2004 (first entry)

XX DE Human NOV1a protein SEQ ID NO:2.

XX KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;

XX KW renal cancer; inflammation; tissue typing.

XX OS Homo sapiens.

XX FN WO2003080856-A2.

XX PD 02-OCT-2003.

XX 19-MAR-2003; 2003WO-US008490.

XX 19-MAR-2002; 2002US-0365491P.

PR 13-SEP-2002; 2002US-0410618P.

XX (CURA-) CURAGEN CORP.

XX PA

XX PI Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
 PI Ooi CS, Anderson DW, Guo X, Giot L, Starling G;
 XX WPI: 2003-876927/81.
 DR N-PSDB; ADE36591.
 XX New polypeptide, useful for preparing a composition for treating or
 PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
 PT or inflammation, or for tissue typing.
 XX Claim 1; SEQ ID NO 2; 239pp; English.
 XX The present invention describes an isolated human NOVX polypeptide, where
 CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
 CC least 95 % identical to it, or a sequence comprising one or more
 CC conservative substitutions in the amino acid sequence. The NOVX
 CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
 CC anti-inflammatory activities, and can be used in gene therapy, and in
 CC vaccines. The NOVX polypeptide is useful for preparing a composition for
 CC treating or preventing a pathology associated with NOVX polypeptide e.g.
 CC renal cancer or inflammation, or for tissue typing. The present sequence
 CC represents human NOV1a from the present invention.
 XX Sequence 339 AA;
 SQ Query Match 100.0%; Score 432; DB 7; Length 339;
 Best Local Similarity 100.0%; Pred. No. 9.5e-36;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPLPRQNHVPVATSPSSPQPAETHPTTLOGAIRREPTSSPLYTTGNDTVTSSDGLW 60
 DB 190 MPLPRQNHVPVATSPSSPQPAETHPTTLOGAIRREPTSSPLYTTGNDTVTSSDGLW 249
 QY 61 NNNQTLFLEHSLTANTTKG 81
 DB 250 NNNQTLFLEHSLTANTTKG 270
 RESULT 8
 ABR58582
 ID ABR58582 standard; protein; 359 AA.
 XX AC ABR58582;
 XX 09-JUL-2003 (first entry)
 XX Human cancer related protein SEQ ID NO:239.
 XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
 KW heart disease; atherosclerosis; endometriosis.
 XX Homo sapiens.
 XX WO2003025138-A2.
 XX 27-MAR-2003.
 XX 17-SEP-2002; 2002WO-US029560.
 XX 17-SEP-2001; 2001US-0323469P.
 PR 20-SEP-2001; 2001US-0323887P.
 PR 13-NOV-2001; 2001US-0350666P.
 PR 08-FEB-2002; 2002US-0355145P.
 PR 08-FEB-2002; 2002US-0355257P.
 PR 12-APR-2002; 2002US-037246P.
 XX (EOS-) EOS BIOTECHNOLOGY INC.
 XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
 PI Zlotnik A;
 XX WPI: 2003-354600/33.

DR N-PSDB; ACC72709.
 XX New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 XX Claim 12; Page 742; 767pp; English.
 XX The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in
 CC ABR58521 to ABR58709. Also described: (1) determining the presence or
 CC absence of a pathological cell in a patient; (2) an expression vector
 CC comprising a nucleic acid molecule described above; (3) a host cell
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 CC of (4); (6) specifically targeting a compound to a pathological cell in a
 CC patient by administering to the patient the antibody above; and (7) a
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 CC therapeutic targets. In particular, the nucleic acid is useful for
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in
 CC drug screening, particularly for identifying agents for treating these
 CC pathologies
 XX Sequence 359 AA;
 SQ Query Match 100.0%; Score 432; DB 6; Length 359;
 Best Local Similarity 100.0%; Pred. No. 1e-35;
 Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MPLPRQNHVPVATSPSSPQPAETHPTTLOGAIRREPTSSPLYTTGNDTVTSSDGLW 60
 DB 210 MPLPRQNHVPVATSPSSPQPAETHPTTLOGAIRREPTSSPLYTTGNDTVTSSDGLW 269
 QY 61 NNNQTLFLEHSLTANTTKG 81
 DB 270 NNNQTLFLEHSLTANTTKG 290
 RESULT 9
 ABR70439
 ID ABR70439 standard; protein; 359 AA.
 XX AC ABR70439;
 XX 22-APR-2003 (first entry)
 XX Amino acid sequence of human TIM-1 allele 1.
 XX T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
 KW TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
 KW myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
 KW allergic T cell response; autoimmune disease.
 XX Homo sapiens.
 XX WO2003002722-A2.
 XX 09-JAN-2003.
 XX 01-JUL-2002; 2002WO-US020890.
 XX 29-JUN-2001; 2001US-0302344P.
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
 XX

[illegible]

CC The present invention describes a method for detecting a bladder cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is 80 % identical to a table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059 encode the human bladder cancer-associated proteins given in ABR48146 to ABR49242). Bladder cancer-associated sequences from the present invention have cytostatic activities, and can be used in antisense gene therapy and in vaccine production. The method can be used for detecting a bladder cancer-associated transcript in a cell from a patient. The method is useful in diagnosing or treating bladder cancer and in screening for compounds that modulate bladder cancer, such as hormones or antibodies. The nucleic acid molecules from the present invention may be used in various screening and diagnostic methods, and for gene therapy, vaccine and/or antisense/inhibition applications

XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 432; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESSDGLW 60
|||
DB 210 MPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESSDGLW 269

QY 61 NNNQTQLFLEHSLLTANTTKG 81
|||
DB 270 NNNQTQLFLEHSLLTANTTKG 290

RESULT 12
AAO26680
ID AAO26680 standard; protein; 359 AA.

XX AAO26680;

XX 20-MAR-2003 (first entry)

DE KIM-1 related protein, SEQ ID No 8.

XX Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.

XX Homo sapiens.

XX WO200298920-A1.

XX 12-DEC-2002.

XX 31-MAY-2002; 2002WO-US017402.

XX 01-JUN-2001; 2001US-0295449P.

XX 04-JUN-2001; 2001US-0295907P.

XX (BIOJ) BIOGEN INC.

XX (GEO) GEN HOSPITAL CORP.

XX Bailly V, Bonventre J;

XX WPI; 2003-156845/15.

XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.

XX Disclosure; Page 40-41; 42pp; English.

XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.

CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a protein relating to the KIM-1
CC protein of the invention

XX Sequence 359 AA;

Query Match 100.0%; Score 432; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESSDGLW 60
|||
DB 210 MPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLSYTTDGNVTVESSDGLW 269

QY 61 NNNQTQLFLEHSLLTANTTKG 81
|||
DB 270 NNNQTQLFLEHSLLTANTTKG 290

RESULT 13

ADE36594
ID ADE36594 standard; protein; 359 AA.

XX ADE36594;

XX 29-JAN-2004 (first entry)

DE Human NOV1b protein SEQ ID NO:4.

XX human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.

XX Homo sapiens.

XX WO2003080856-A2.

XX 02-OCT-2003.

XX 19-MAR-2003; 2003WO-US008490.

XX 19-MAR-2002; 2002US-0365491P.

XX 13-SEP-2002; 2002US-0410618P.

XX (CURA-) CURAGEN CORP.

XX Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;

PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;

XX WPI; 2003-876927/81.

XX N-PSDB; ADE36593.

XX New polypeptide, useful for preparing a composition for treating or

PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer

PT or inflammation, or for tissue typing.

XX Claim 1; SEQ ID NO 4; 239pp; English.

XX The present invention describes an isolated human NOVX polypeptide, where
CC X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to it, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOV1b from the present invention.

XX Sequence 359 AA;

Query Match 100.0%; Score 432; DB 7; Length 359;

Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHVPATSSSPQPAETHPTTLOGAIRRPTSSPLYSTTDCGNDVTTESSDGLW 60
DB 210 MPLPRQNHVPATSSSPQPAETHPTTLOGAIRRPTSSPLYSTTDCGNDVTTESSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
DB 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 14
ADN38984
ID ADN38984 standard; protein; 359 AA.
XX
AC ADN38984;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:302.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerrary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI: 2003-468649/44.
DR N-PSDB; ADN38983.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO 302; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC

CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 432; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 1e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHVPATSSSPQPAETHPTTLOGAIRRPTSSPLYSTTDCGNDVTTESSDGLW 60
DB 210 MPLPRQNHVPATSSSPQPAETHPTTLOGAIRRPTSSPLYSTTDCGNDVTTESSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
DB 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 15
ADQ76690
ID ADQ76690 standard; protein; 359 AA.
XX
AC ADQ76690;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human kidney injury molecule-1 (KIM-1).
XX
KW Human; KIM-1; kidney injury module-1; immunosuppressive; protozoacide;
KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW antiallergic; antiinflammatory; vasotropic; nephropic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT Modified-site /label= Signal_peptide
FT Modified-site 65..67
FT Modified-site /note= "N-glycosylated"
FT Modified-site 258..260
FT Modified-site /note= "N-glycosylated"
FT Modified-site 272..274
FT Modified-site /note= "N-glycosylated"
FT Modified-site 286..288
FT Domain /note= "N-glycosylated"
FT Domain 290..311
FT Region /note= "Transmembrane domain"
FT Region 324..359
FT /note= "C-terminal region unique to this splice variant"

WO2004060041-A2.
XX
XX 22-JUL-2004.
XX
XX 29-DEC-2003; 2003WO-US041294.
XX
XX 30-DEC-2002; 2002US-0436934P.
XX
XX (BIOG-) BIOGEN IDEC MA INC.
XX

PI Rennert PD;
XX WPI; 2004-534277/51.
XX
XX
PT Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
PT Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
PT activation, or inhibiting production of a subset of antibodies against
PT one or more antigens.
XX
XX
PS Disclosure; SEQ ID NO 1; 68pp; English.
XX
XX The present sequence is that of human kidney injury molecule-1 (KIM-1), a
CC type I cell membrane glycoprotein and member of the T cell immunoglobulin
CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids
CC ADQ76696 or 359 amino acids, depending on splice variation. It has been
CC discovered that treatment of a mammal with a KIM-1 antagonist alters the
CC interaction of T cells and other immune system cells, e.g. dendritic
CC cells, monocytes, macrophages and B cells, and thereby strongly
CC suppresses an IgG response to an antigen. Such treatment also eliminates
CC IgG1 production by memory B cells in response to subsequent challenge
CC with the antigen. Blockage of the binding of KIM-1 to its receptor
CC reduces secretion of interferon-gamma by immune cells engaged in an
CC antigen response in the mixed lymphocyte response assay. Based on these
CC discoveries, the invention provides methods for therapeutically
CC modulating immune function in autoimmune diseases and other disorders of
CC the mammalian immune system. A method is claimed for inhibiting
CC signalling between a T cell and a second cell, e.g. an antigen-presenting
CC cell, in a mammal. The method involves identifying a mammal with an
CC immune disease or disorder, or one preparing to receive a tissue graft,
CC and administering a KIM-1 antagonist, especially a polypeptide
CC comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a
CC KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding
CC fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a
CC soluble polypeptide, which can include a KIM-1 mucin domain in addition
CC to the KIM Ig domain, and may include a heterologous moiety such as an Fc
CC moiety. The antagonist may be conjugated to a polymer such as
CC polyethylene glycol. The KIM-1 antagonist is used in methods of:
CC inhibiting activation of a B cell in a mammal; inhibiting disease relapse
CC in an autoimmune disease; inhibiting epitope spreading in an autoimmune
CC disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis,
CC autoimmune haemolytic anaemia, Chagas disease, Grave's disease,
CC idiopathic thrombocytopenia purpura, Wegener's granulomatosis,
CC polyarteritis nodosa, rapidly progressive crescentic glomerulonephritis,
CC graft-versus-host disease, or systemic lupus nephritis) or inflammatory
CC disease or disorder (e.g. inflammatory bowel diseases such as Crohn's
CC disease, ulcerative colitis, and ileitis); and inhibiting secretion of
CC IFN-gamma by lymphocytes.
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 432; DB 8; Length 359;
Best Local Similarity 100.0%; Pred. NO. 1e-35;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSGYTTDGNVTVESDGLW 60
DB 210 MPLPRQNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSGYTTDGNVTVESDGLW 269

QY 61 NNNQTQLFLEHSLLTANTTKG 81
DB 270 NNNQTQLFLEHSLLTANTTKG 290

Search completed: June 29, 2005, 09:09:06
Job time : 59.6071 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:58:20 ; Search time 50.7642 Seconds
(without alignments)
613.589 Million cell updates/sec

Title: US-10-718-321-6
Perfect score: 432
Sequence: 1 MFLPRQNHPEVATSSSSPOP.....NNQTLFLEHSLLTANTTKG 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA.*
- 1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/prodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/prodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/prodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/prodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/prodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/prodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/prodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/prodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/prodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/prodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/prodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/prodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/prodata/1/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/prodata/1/pubpaa/US10F_PUBCOMB.pep.*
 - 19: /cgn2_6/prodata/1/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/prodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/prodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/prodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	432	100.0	81	17 US-10-718-321-6	Sequence 6, Appli
2	432	100.0	334	17 US-10-655-506-7	Sequence 7, Appli
3	432	100.0	334	17 US-10-718-321-7	Sequence 7, Appli
4	432	100.0	339	17 US-10-391-939A-2	Sequence 2, Appli
5	432	100.0	339	17 US-10-391-939A-28	Sequence 28, Appli
6	432	100.0	359	14 US-10-391-939A-32	Sequence 17, Appl
7	432	100.0	359	14 US-10-188-012-17	Sequence 19, Appl
8	432	100.0	359	15 US-10-188-012-19	Sequence 302, App
9	432	100.0	359	15 US-10-295-027-302	Sequence 64, Appl
10	432	100.0	359	15 US-10-188-832-64	Sequence 4, Appli
11	432	100.0	359	17 US-10-391-939A-4	Sequence 17, Appl

12	432	100.0	359	17 US-10-663-497-19	Sequence 19, Appl
13	432	100.0	359	17 US-10-718-321-8	Sequence 8, Appli
14	432	100.0	359	17 US-10-847-918-25	Sequence 25, Appl
15	432	100.0	364	14 US-10-188-012-25	Sequence 25, Appl
16	432	100.0	364	14 US-10-663-497-25	Sequence 25, Appl
17	432	100.0	365	14 US-10-188-012-21	Sequence 21, Appl
18	432	100.0	365	14 US-10-663-497-21	Sequence 21, Appl
19	427	98.8	364	17 US-10-188-012-27	Sequence 27, Appl
20	427	98.8	364	17 US-10-663-497-27	Sequence 27, Appl
21	423	97.9	359	14 US-10-188-012-23	Sequence 23, Appl
22	423	97.9	359	17 US-10-663-497-23	Sequence 23, Appl
23	396	91.7	263	17 US-10-391-939A-6	Sequence 6, Appli
24	386	91.7	263	17 US-10-391-939A-32	Sequence 32, Appl
25	396	91.7	263	17 US-10-391-939A-39	Sequence 39, Appl
26	396	91.7	263	17 US-10-805-177-50	Sequence 50, Appl
27	99	22.9	18	17 US-10-718-321-1	Sequence 1, Appli
28	89	20.6	307	17 US-10-655-506-3	Sequence 3, Appli
29	78.5	18.2	141	15 US-10-425-114-52586	Sequence 52586, A
30	78.5	18.2	144	16 US-10-425-115-245392	Sequence 245392, A
31	77.5	17.9	418	16 US-10-437-963-182939	Sequence 182939, A
32	77	17.8	440	17 US-10-732-923-6569	Sequence 6569, Ap
33	77	17.8	441	17 US-10-732-923-6588	Sequence 6588, Ap
34	77	17.8	483	16 US-10-437-963-157601	Sequence 157601, A
35	76.5	17.7	282	14 US-10-188-012-3	Sequence 3, Appli
36	76.5	17.7	282	17 US-10-663-497-3	Sequence 3, Appli
37	76	17.6	647	16 US-10-821-273-6	Sequence 6, Appli
38	76	17.6	893	15 US-10-334-143-27	Sequence 27, Appl
39	76	17.6	2115	16 US-10-437-963-195579	Sequence 195579, A
40	75.5	17.5	235	16 US-10-425-115-331823	Sequence 331823, A
41	75.5	17.5	1588	16 US-10-437-963-189741	Sequence 189741, A
42	75	17.4	8360	14 US-10-132-134-34	Sequence 34, Appl
43	73	16.9	127	15 US-10-424-599-250789	Sequence 250789, A
44	73	16.9	579	16 US-10-437-963-151894	Sequence 151894, A
45	72.5	16.8	302	15 US-10-225-066A-778	Sequence 778, App

ALIGNMENTS

RESULT 1
US-10-718-321-6
; Sequence 6, Application US/10718321
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; TITLE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-6

Query Match 100.0%; Score 432; DB 17; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.4e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFLPRQNHPEVATSSSSPOPASTHTTLOGATRRRPTSSPLSYTTGNDTVTSSDGLW 60
|||||


```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-188-012-17

Query Match          100.0%; Score 432; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDGNDTVTSSDGLW 60
   |||||||
Db 210 MPLPRNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDGNDTVTSSDGLW 269
   |||||||

QY 61 NNNQTQLFLEHSLLTANTTKG 81
   |||||||
Db 270 NNNQTQLFLEHSLLTANTTKG 290
   |||||||

RESULT 7
US-10-188-012-19
; Sequence 19, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)..(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-188-012-19

Query Match          100.0%; Score 432; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDGNDTVTSSDGLW 60
   |||||||
Db 210 MPLPRNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDGNDTVTSSDGLW 269
   |||||||

QY 61 NNNQTQLFLEHSLLTANTTKG 81
   |||||||
Db 270 NNNQTQLFLEHSLLTANTTKG 290
   |||||||

RESULT 8
US-10-295-027-302
; Sequence 302, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha

```

```
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-302

Query Match      100.0%; Score 432; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNEHPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNNDVTSSDGLW 60
Db 210 MPLPRQNEHPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNNDVTSSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
Db 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 9
US-10-188-832-64
; Sequence 64, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Nataasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
```

```
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-64

Query Match      100.0%; Score 432; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNEHPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNNDVTSSDGLW 60
Db 210 MPLPRQNEHPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNNDVTSSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
Db 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 10
US-10-391-939A-4
; Sequence 4, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Giot, Ioic ;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Mesri, Mehdi ;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca ;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glennda ;
; APPLICANT: Starling, Gary ;
; APPLICANT: Tse, Kam-Pai
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-048
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-4

Query Match      100.0%; Score 432; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNEHPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNNDVTSSDGLW 60
Db 210 MPLPRQNEHPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNNDVTSSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
Db 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 11
```

US-10-663-497-17
; Sequence 17, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; PRIOR FILING DATE: 2003-09-15
; PRIOR FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-663-497-17

Query Match 100.0%; Score 432; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHPEVATSPSSQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTTESSDGLW 60
DB 210 MPLPRQNHPEVATSPSSQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTTESSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
DB 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 12
US-10-663-497-19
; Sequence 19, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR FILING DATE: 2001-06-29
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-663-497-19

Query Match 100.0%; Score 432; DB 17; Length 359;

Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHPEVATSPSSQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTTESSDGLW 60
DB 210 MPLPRQNHPEVATSPSSQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTTESSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
DB 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 13
US-10-718-321-8
; Sequence 8, Application US/10718321
; Publication No. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; TITLE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-8

Query Match 100.0%; Score 432; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHPEVATSPSSQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTTESSDGLW 60
DB 210 MPLPRQNHPEVATSPSSQPAETHPTTLOGAIRREPTSSPLSYTTDGNDDVTTESSDGLW 269

QY 61 NNNQTLFLEHSLLTANTTKG 81
DB 270 NNNQTLFLEHSLLTANTTKG 290

RESULT 14
US-10-847-918-25
; Sequence 25, Application US/10847918
; Publication No. US20050119210A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Be, Xiaobing
; APPLICANT: Liu, Wei
; APPLICANT: Slonim, Donna
; APPLICANT: Howes, Steve
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating Cancers
; FILE REFERENCE: 031896-026000 (AM101264)
; CURRENT APPLICATION NUMBER: US/10/847,918
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US 60/471,729
; PRIOR FILING DATE: 2003-05-20
; NUMBER OF SEQ ID NOS: 14937
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 359

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-847-918-25

Query Match      100.0%; Score 432; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 8.3e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MPLPRONHEPVATSPSSPOPAETHPTTLOGAIRREPTSSPLSYTTDGNDTVTESSDGLW 60
Db      210 MPLPRONHEPVATSPSSPOPAETHPTTLOGAIRREPTSSPLSYTTDGNDTVTESSDGLW 269

QY      61  NNNQTOLFLEHSLLTANTTKG 81
Db      270 NNNQTOLFLEHSLLTANTTKG 290
```

```
RESULT 15
US-10-188-012-25
; Sequence 25, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; TITLE OF INVENTION: Use Thereof
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 364
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(364)
; OTHER INFORMATION: TIM-1 allele 5
US-10-188-012-25
```

```
Query Match      100.0%; Score 432; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 8.5e-36;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MPLPRONHEPVATSPSSPOPAETHPTTLOGAIRREPTSSPLSYTTDGNDTVTESSDGLW 60
Db      215 MPLPRONHEPVATSPSSPOPAETHPTTLOGAIRREPTSSPLSYTTDGNDTVTESSDGLW 274

QY      61  NNNQTOLFLEHSLLTANTTKG 81
Db      275 NNNQTOLFLEHSLLTANTTKG 295
```

Search completed: June 29, 2005, 09:33:00
Job time : 51.7642 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 7.32024 Seconds
(without alignments)
1064.659 Million cell updates/sec

Title: US-10-718-321-6
Perfect score: 432
Sequence: 1 MFLPRQNHVPVATSPSPQP.....NNQTQLFLEHSLLTANTTKG 81
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	348	80.6	451	2 S71754	cellular hepatitis
2	79.5	18.4	1007	2 T24643	hypothetical prote
3	76	17.6	1006	2 T42731	atrophin-1 related
4	72.5	16.8	1082	2 T43204	chromatin structur
5	72	16.7	235	2 FC2022	mucin like protein
6	72	16.7	1402	2 I46707	translation initia
7	72	16.7	1777	2 T34369	hypothetical prote
8	71	16.4	1131	2 T15787	hypothetical prote
9	70.5	16.3	644	2 S39356	transcription fact
10	70.5	16.3	648	2 T20144	hypothetical prote
11	70.5	16.3	1095	2 FC1114	SKCDC25 protein -
12	70	16.2	321	2 T45053	hypothetical prote
13	70	16.2	654	2 S69673	SAC7 protein - yea
14	69.5	16.1	479	2 G71957	probable outer mem
15	69.5	16.1	542	2 I39540	chitinase (EC 3.2.
16	69.5	16.1	574	2 T05964	probable low-affin
17	69	16.0	602	2 A45769	acetylcholine rece
18	69	16.0	1166	2 T28680	fibrogen-binding
19	69	16.0	1317	2 A54831	nuclear pore compl
20	68.5	15.9	213	2 T01715	hypothetical prote
21	68.5	15.9	285	2 G85016	probable myb-relat
22	68.5	15.9	681	2 A45705	type 1 transmembra
23	68.5	15.9	1042	2 E85968	evolved beta-D-gal
24	68.5	15.9	1042	2 F91123	evolved beta-D-gal
25	68.5	15.9	1370	2 T19188	hypothetical prote
26	68.5	15.9	2241	2 T20971	hypothetical prote
27	68.5	15.9	2261	2 T20978	hypothetical prote
28	68	15.7	310	2 T45873	hypothetical prote
29	68	15.7	405	2 H89930	protein RllG11.14

30	68	15.7	620	2 A70525	hypothetical prote
31	68	15.7	1056	2 A53767	mucin MUC5B, trach
32	68	15.7	2018	2 T34274	hypothetical prote
33	67.5	15.6	137	2 G96017	conserved hypotet
34	67.5	15.6	1419	2 T30531	agglutinin-like ad
35	67	15.5	447	2 T18447	hrpW protein - Erw
36	67	15.5	1015	2 JC6552	DNA topoisomerase
37	67	15.5	1186	2 T19050	hypothetical prote
38	67	15.5	1560	2 T42727	proliferation pote
39	66.5	15.4	322	2 T23891	hypothetical prote
40	66.5	15.4	398	2 S50507	excision repair pr
41	66.5	15.4	575	2 S39484	DNA-binding protei
42	66.5	15.4	881	2 T28013	hypothetical prote
43	66.5	15.4	1589	2 C44766	defective chorion-
44	66	15.3	437	2 S15704	transforming prote
45	66	15.3	575	2 JC7794	lammer kinase homo

ALIGNMENTS

RESULT 1

S71754
cellular hepatitis A receptor HAVcr-1 precursor - green monkey
N;Alternate names: surface glycoprotein
C:Species: Cercopithecus aethiops (green monkey, grivet)
C>Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 09-Jul-2004
C:Accession: S71754
R:Kaplan, G.; Totsuka, A.; Thompson, P.; Akatsuka, T.; Moritsugu, Y.; Feinstone, S.M.
EMBO J. 15, 4282-4296, 1996
A>Title: Identification of a surface glycoprotein on African green monkey kidney cells as
A:Reference number: S71754; MUID:97015129; PMID:8861957
A:Accession: S71754
A:Molecule type: mRNA
A:Residues: 1-451 <KAP>
A:Cross-references: UNIPROT:Q95144; EMBL:X98252; NID:gl526573; PID:e247449; PID:gl526574
A:Experimental source: kidney
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-451/Product: cellular hepatitis A receptor HAVcr-1 #status predicted <MAT>

Query Match 80.6%; Score 348; DB 2; Length 451;
Best Local Similarity 82.5%; Pred. No. 1.5e-26;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

Qy 2 PLPRQNHVPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYVTDPGNDVTVESSDGLWN 61
Db 291 PLPMQNHVPVATSPSPQPAETHPTTLOGAIRREPTSSPLSYVTDPGNDVTVESSDGLWN 350

Qy 62 NNQTQLFLEHSLLTANTTKG 81
Db 351 NNQTQLSPHSPQWNTTEG 370

RESULT 2

T24643
hypothetical protein T07C12.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T24643
R:McMurray, A.
A:Reference number: Z19916
submitted to the EMBL Data Library, June 1996
A:Accession: T24643
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1007 <WIL>
A:Cross-references: UNIPROT:Q22286; EMBL:Z73976; PIDN:CAA98287.1; GSPDB:GN00023; CESP:T07C12.8
A:Experimental source: clone T07C12
C:Genetics:
A:Gene: CESP:T07C12.8
A:Map position: 5
A:Introns: 137/1; 178/2; 275/1; 295/3; 336/2; 390/1; 747/1; 846/3; 890/1; 953/2

```
Query Match      18.4%; Score 79.5; DB 2; Length 1007;
Best Local Similarity 26.1%; Pred. No. 8.1;
Matches 23; Conservative 11; Mismatches 31; Indels 23; Gaps 3;

QY 4 PRQNHVPATSPSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNQDVTWSSD----- 57
DB 540 PSENLIPGTLPSNPAPID-----VRRSPVASS-HFYTNQPTLVTDNDQENQKTL 589
QY 58 -----GLWNNNQTLFLEHSLLTANT 78
DB 590 VDLAKKFGLEWESSPYLTTHPIARNT 617

RESULT 3
T42731
atrophin-1 related protein - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T42731
R:Khan, F.A.; Margolis, R.L.; Loev, S.L.; Sharp, A.H.; Li, S.H.; Ross, C.A.
submitted to the EMBL Data Library, December 1995
A:Description: cDNA sequence and expression of an atrophin-1 (DRPLA disease gene) related
A:Reference number: Z22250
A:Accession: T42731
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1006 <KHA>
A:Cross-references: UNIPROT:Q62901; EMBL:U44091; NID:g1297310; PID:g1209103; PIDN:AAA989
C:Genetics:
A:Gene: ARP

Query Match      17.6%; Score 76; DB 2; Length 1006;
Best Local Similarity 33.9%; Pred. No. 18;
Matches 20; Conservative 9; Mismatches 24; Indels 6; Gaps 3;

QY 2 PLPRQNHPP--VATSPSPQPAETHPTTLQGAIRREPTSSPL--YSYTTDGNQDVTWESS 56
DB 434 PLPSSPAQPGGLTQSQLPPPAASHPTT--GGHLQVPSQSPFPQHPFVGGPPPTPPS 490

RESULT 4
T42204
chromatin structural protein homolog Supt5hp - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T42204
R:Chiang, P.W.; Stubbs, L.; Zhang, L.; Kurmit, D.M.
Genomics 47, 426-428, 1998
A:Title: Isolation of murine SPT5 homologue: completion of the isolation and characteriz
A:Reference number: Z22072; MUID:98149995; PMID:9480761
A:Accession: T42204
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1082 <CHI>
A:Cross-references: UNIPROT:O55201; EMBL:U88539; NID:g2754751; PID:g2754752; PIDN:AA400
C:Genetics:
A:Gene: Supt5h

Query Match      16.8%; Score 72.5; DB 2; Length 1082;
Best Local Similarity 28.6%; Pred. No. 43;
Matches 26; Conservative 11; Mismatches 33; Indels 21; Gaps 4;

QY 2 PLPRQNHVPATSPSPQ-----PASTHTPTTLQGAIRREPTSSPL-----YS 43
DB 900 PSPQSYHQ-VAPSPAGYQNTHTSPASVHTPTSPMAYQASPSFVGYSPTWPGAPSPGGYN 958
QY 44 YTTDGNQDVTWESSDGLWNNNQTLFLEHSL 74
DB 959 PHTPGSGIEQNSSD--WVTTDIQVKVRDYL 987

RESULT 5
```

```
PC2022
mucin like protein Muc2 precursor - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 16-Aug-2004
C:Accession: PC2022
R:Hansson, G.C.; Baeckstroem, D.; Carlstedt, I.; Klinga-Levan, K.
Biochem. Biophys. Res. Commun. 198, 181-190, 1994
A:Title: Molecular cloning of a cDNA coding for a region of an apoprotein from the 'inso1
A:Reference number: PC2022; MUID:94121629; PMID:8292021
A:Accession: PC2022
A:Molecule type: mRNA
A:Residues: 1-235 <HAN>
A:Cross-references: UNIPROT:Q63349; GB:Z29072; NID:g435534; PIDN:CAA82313.1; PID:g435535
A:Experimental source: intestine
C:Genetics:
A:Gene: muc2
A:Map position: 1
C:Superfamily: von Willebrand factor type A repeat homology; von Willebrand factor type (c
C:Keywords: glycoprotein
F:1-53/Region: cysteine-rich
F:54-235/Region: serine/threonine-rich

Query Match      16.7%; Score 72; DB 2; Length 235;
Best Local Similarity 30.4%; Pred. No. 7.4;
Matches 17; Conservative 7; Mismatches 32; Indels 0; Gaps 0;

QY 10 PVATSPSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNQDVTWESSDGLWNNQOT 65
DB 167 PIVTETSPRSTTTQTPTSTVTTTGTPTTTTGTGKTGTTTPTSTVTTSTPTSTTTQT 222

RESULT 6
I46707
translation initiation factor eIF4-gamma - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 09-Jul-2004
C:Accession: I46707
R:Yan, R.; He, W.; Rhoads, R.E.
J. Biol. Chem. 268, 19200-19203, 1993
A:Title: Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of
A:Reference number: I46707; MUID:93374895; PMID:8396129
A:Accession: I46707
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1402 <VAN>
A:Cross-references: UNIPROT:P41110; GB:L22090; NID:g404774; PIDN:AAA31242.1; PID:g404775

Query Match      16.7%; Score 72; DB 2; Length 1402;
Best Local Similarity 35.3%; Pred. No. 67;
Matches 18; Conservative 5; Mismatches 22; Indels 6; Gaps 1;

QY 3 LPRQNHVPATSPSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNQDVT 53
DB 51 LPGAHPSPSPSPSPSPSPPPIL-----EPGSEPNLAVLSLPGDGMT 95

RESULT 7
T34369
hypothetical protein T19D12.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T34369
R:Pavello, A.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid T19D12.
A:Reference number: Z21513
A:Accession: T34369
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1777 <FAV>
A:Cross-references: UNIPROT:Q22579; EMBL:U41263; PIDN:AAC24428.1; GSPDB:GN000020; CESP:T11
A:Experimental source: strain Bristol N2; clone T19D12
C:Genetics:
```

```

Qy 40 -----PLYSYTTD---GNDVTVTESDGLWNNQTQLF 68
Db 221 DLGAAVAASAYGWMNTAYSGGLPARSQFFYAQVASYDYG-N-AVGMSSSAAWFSHOERLY 278

RESULT 10
T20144
hypothetical protein C52A11.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T20144
R:Sulston, J.
submitted to the EMBL Data Library, November 1994
A:Reference number: Z19229
A:Accession: T20144
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-648 <WIL>
A:Cross-references: EMBL:Z46792; PIDN:CAA86766.1; GSPDB:GN00020; CESP:C52A11.1
A:Experimental source: clone C52A11
C:Genetics:
A:Gene: CESP:C52A11.1
A:Map position: 2
A:Introns: 87/3; 237/3; 272/2; 456/1; 565/1; 622/1

Query Match 16.3%; Score 70.5; DB 2; Length 648;
Best Local Similarity 32.6%; Pred. No. 36;
Matches 28; Conservative 10; Mismatches 19; Indels 29; Gaps 6;

Qy 8 HEP-----VATSPSSPP-----AETH-----PTTLOGAIRREPTSSPLYS 43
Db 156 HEKPEPLKVKDTPILPQPVRRTSKEFVTPVKAEHLDAEPTSMPPK-APEPTAPLR 214

Qy 44 YTTDGNDF-VTESSDGLWN--NNQT 65
Db 215 LSTDSDTSFSGHSSKDLFSPTNPQT 240

RESULT 11
PC1114
SKDC25 protein - yeast (Saccharomyces kluyveri) (fragment)
C:Species: Saccharomyces kluyveri
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: PC1114
R:Prigozy, T.; Gonzales, E.; Broek, D.
Gene 117, 67-72, 1992
A:Title: Identification and analysis of a DNA fragment from Saccharomyces kluyveri that
A:Reference number: PC1114; MUID:92354938; PMID:1644315
A:Accession: PC1114
A:Molecule type: DNA
A:Residues: 1-1095 <PRI>
A:Cross-references: UNIPROT:Q02342; GB:M82964; NID:G171186; PIDN:AAA34479.1; PID:G171187
C:Genetics:
A:Gene: SKDC25
C:Keywords: transmembrane protein
F:808-1049/Domain: CDC25-type guanine nucleotide exchange activator homology <SOS>

Query Match 16.3%; Score 70.5; DB 2; Length 1095;
Best Local Similarity 23.1%; Pred. No. 69;
Matches 15; Conservative 19; Mismatches 28; Indels 3; Gaps 2;

Qy 4 PRQNHPEV-ATSPSPQPAETHPTTLOGAIRREPTSSPLYSYTTDGNVTESDGLWNN 62
Db 104 PRDSKVSITNTTIPDFSRSSASTIPSLFNOESPARPINFSSTTDTLMFSD--WQS 161

Qy 63 NOTQL 67
Db 162 NSPNL 166

RESULT 12
T45053
hypothetical protein v1068a [imported] - Caenorhabditis elegans

```

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 29.8703 Seconds
(without alignments)
543.815 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225

Sequence: 1 PLYSTTDGNDVTSSDGL.....NNQTQLFLESLTANTYKG 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003Bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	225	100.0	81	AAO26678	Human KIM
2	225	100.0	298	ADQ76693	Human KIM
3	225	100.0	334	AAW38336	Human kid
4	225	100.0	334	AAO26679	KIM-1 rel
5	225	100.0	334	ADQ76696	Human kid
6	225	100.0	339	ADQ76696	Human kid
7	225	100.0	339	ADQ76696	Human kid
8	225	100.0	359	ADQ76696	Human kid
9	225	100.0	359	ADQ76696	Human kid
10	225	100.0	359	ADQ76696	Human kid
11	225	100.0	359	ADQ76696	Human kid
12	225	100.0	359	ADQ76696	Human kid
13	225	100.0	359	ADQ76696	Human kid
14	225	100.0	359	ADQ76696	Human kid
15	225	100.0	359	ADQ76696	Human kid
16	225	100.0	364	ADQ76696	Human kid
17	225	100.0	365	ADQ76696	Human kid
18	225	100.0	518	ADQ76696	Human KIM
19	220	97.8	359	ADQ76696	Human kid
20	220	97.8	359	ADQ76696	Human kid
21	189	84.0	263	ADQ76696	Human kid
22	189	84.0	263	ADQ76696	Human kid
23	189	84.0	263	ADQ76696	Human kid
24	178	79.1	451	ADQ76696	Human kid
25	101	44.9	18	AAO26686	Human KIM

26	99	44.0	18	6	AAO26687	Human KIM
27	99	44.0	18	6	AAO26673	Monoclonal
28	90	40.0	18	6	AAO26688	Human KIM
29	82	36.4	18	6	AAO26685	Human KIM
30	60	26.7	602	2	AAW51705	Neurotrop
31	60	26.7	602	2	AAW32900	Chicken n
32	60	26.7	606	4	AAW78898	C. glutam
33	60	26.7	1070	4	AAW48099	Amino aci
34	60	26.7	1070	4	AAW48099	Chicken n
35	58.5	26.0	774	8	AAW48099	Chicken n
36	58.5	26.0	775	8	AAW48099	Chicken n
37	58	25.8	981	5	AAW48099	Venezuela
38	57	25.3	307	2	AAW38334	Rat kidne
39	57	25.3	679	8	AAW38334	Bacterial
40	57	25.3	2399	6	AAW38334	Protein e
41	56.5	25.1	1658	5	AAW38334	Protein e
42	56	24.9	219	3	AAW38334	Arabidops
43	56	24.9	249	3	AAW38334	Arabidops
44	56	24.9	249	3	AAW38334	Arabidops
45	56	24.9	268	3	AAW38334	Arabidops

ALIGNMENTS

RESULT 1
AAO26678
ID AAO26678 standard; protein; 81 AA.
XX AAO26678;
AC AAO26678;
XX 20-MAR-2003 (first entry)
XX Human KIM-1 mucin domain protein, SEQ ID No 6.
DE Human KIM-1 mucin domain protein, SEQ ID No 6.
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX Homo sapiens.
XX WO200298920-A1.
XX 12-DEC-2002.
XX 31-MAY-2002; 2002WO-US017402.
XX 01-JUN-2001; 2001US-0295449P.
XX 04-JUN-2001; 2001US-0295907P.
XX (BIO) BIOGEN INC.
XX (GEO) GEN HOSPITAL CORP.
XX Bailly V, Bonventre J;
XX WPI; 2003-156845/15.
XX New antibody, antibody derivative or antigen-binding polypeptide that
XX inhibits proteolytic release of a soluble kidney injury molecule-1
XX polypeptide, useful for treating or preventing renal disease or injury,
XX e.g. renal cancer.
XX Disclosure; Fig 1A; 42pp; English.
XX The invention relates to a novel antibody, antibody derivative or antigen
XX -binding polypeptide that inhibits proteolytic release of a soluble
XX kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
XX The antibody, antibody derivative or antigen-binding polypeptide is
XX useful for treating or preventing renal disease or injury, e.g. renal
XX cancer. The antibody is also useful for inhibiting shedding of the KIM-1
XX polypeptide. This sequence represents a human KIM-1 mucin domain protein
XX of the invention

SQ Sequence 81 AA;
Query Match 100.0%; Score 225; DB 6; Length 81;
Best Local Similarity 100.0%; Pred. No. 3.3e-21;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PLYSYTDDGNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
DB 40 PLYSYTDDGNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 81
RESULT 2
ADQ76693
ID ADQ76693 standard; protein; 298 AA.
XX AC ADQ76693;
XX 07-OCT-2004 (first entry)
XX Human KIM-1 extracellular domain-6xHis fusion.
XX Human; KIM-1; kidney injury module-1; immunosuppressive; protozoacide;
KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW antiallergic; antiinflammatory; vasotropic; nephrotropic.
XX OS Homo sapiens.
OS Synthetic.
XX FH Key Location/Qualifiers
FT Peptide 1..20 /label= Signal_peptide
XX WO2004060041-A2.
XX 22-JUL-2004.
XX 29-DEC-2003; 2003WO-US041294.
XX 30-DEC-2002; 2002US-0436934P.
XX (BIOG-) BIOGEN IDEC MA INC.
XX Rennert PD;
XX WPI; 2004-534277/51.
XX Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
PT Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
PT activation, or inhibiting production of a subset of antibodies against
PT one or more antigens.
XX Example 3; Page 33; 68pp; English.
XX The present sequence is that of a polypeptide comprising the
CC extracellular domain (residues 1-290) of human kidney injury molecule-1
CC (KIM-1) ADQ76690 fused to a C-terminal 6xHis tag peptide. A plasmid
CC comprising DNA encoding this sequence under control of a CMV promoter was
CC used for transient constitutive expression of the polypeptide in
CC mammalian cells in an example from the invention. The invention provides
CC methods for therapeutically modulating immune function in autoimmune
CC diseases and other disorders of the mammalian immune system. A method is
CC claimed for inhibiting signalling between a T cell and a second cell,
CC e.g. an antigen-presenting cell, in a mammal. The method involves
CC identifying a mammal with an immune disease or disorder, or one preparing
CC to receive a tissue graft, and administering a KIM-1 antagonist,
CC especially a polypeptide comprising: a KIM-1 Ig domain, and lacking a
CC transmembrane domain and a KIM-1 cytoplasmic domain; an anti-KIM-1
CC antibody; or an antigen-binding fragment of an anti-KIM-1 antibody.
CC Preferably, the KIM-1 antagonist is a soluble polypeptide, which can
CC include a KIM-1 mucin domain in addition to the KIM Ig domain, and may
CC include a heterologous moiety such as an Fc moiety. The antagonist may be
CC conjugated to a polymer such as polyethylene glycol. The KIM-1 antagonist
CC is used in methods of: inhibiting activation of a B cell in a mammal;

CC inhibiting disease relapse in an autoimmune disease; inhibiting epitope
CC spreading in an autoimmune disease; treating a Th2 cell-mediated disease
CC (e.g. myasthenia gravis, autoimmune haemolytic anaemia, Chagas disease,
CC Grave's disease, idiopathic thrombocytopenia purpura, Wegener's
CC granulomatosis, polyarteritis nodosa, rapidly progressive crescentic
CC glomerulonephritis, graft-versus-host disease, or systemic lupus
CC nephritis) or inflammatory disease or disorder (e.g. inflammatory bowel
CC diseases such as Crohn's disease, ulcerative colitis, and ileitis); and
CC inhibiting secretion of IFN-gamma by lymphocytes.
XX XX Sequence 298 AA;

Query Match 100.0%; Score 225; DB 8; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.6e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTDDGNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
DB 249 PLYSYTDDGNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 3
AAW38336
ID AAW38336 standard; protein; 334 AA.
XX AC AAW38336;
XX 21-MAY-1998 (first entry)
XX Human kidney injury related molecule (KIM).
XX Kidney injury related molecule; KIM; human; renal disease; injury;
KW nephritis; tissue regeneration; therapy; monoclonal antibody.
XX OS Homo sapiens.
XX WO9744460-A1.
XX 27-NOV-1997.
XX 23-MAY-1997; 97WO-US009303.
XX 24-MAY-1996; 96US-0018228P.
XX 23-AUG-1996; 96US-0023442P.
XX (BIOJ) BIOGEN INC.
XX Sanicola-Nadel M, Bonventre JV, Hession CA, Ichimura T, Wei H;
PI Cate RL;
XX WPI; 1998-018514/02.

DNA encoding kidney injury related molecule - which is upregulated in
injured or regenerating tissue, useful to promote growth of new tissue
and survival of damaged tissue.

Claim 9; Page 46-47; 68pp; English.

This protein, designated kidney injury related molecule (KIM), is up-
regulated in injured or regenerating tissue. Its amino acid sequence was
deduced from a clone (see AAT96035) obtained from a human embryonic liver
library. A 572-amino acid (see AAW38335) and a 307-amino acid rat KIM
(see AAW38334) are also claimed. Recombinant KIM polypeptides can be
expressed in prokaryotic and eukaryotic host cells using a claimed
process. Soluble variants fused to a toxin, imageable compound or
radionuclide, and IgG fusion proteins are also claimed. KIM, or an
agonist, can be used to treat renal disease and to promote the growth of new
tissue or the survival of damaged tissue, generally in conditions where
the binding of specific ligand to KIM stimulates cell growth, maintains
cellular differentiation or reduces apoptosis, e.g. in cases of renal
failure, nephritis, kidney transplants, toxic or hypoxic injury. A
monoclonal antibody specific for KIM can be used to treat renal disease,
e.g. where binding of KIM to ligand results in neoplasia, loss of

CC cellular function, susceptibility to apoptosis or promotion of
CC inflammation, deliver imaging agents to KIM expressing cells in vivo or
CC in vitro and measure KIM concentration by immunoassay.
CC Damage/regeneration of renal cells can be determined by measuring KIM,
CC particularly to diagnose or monitor the progress of disease or therapy.
CC KIM-expressing tumour cells can be inhibited by treatment with a fusion
CC protein comprising KIM ligand or Mab with a toxin or radionuclide, and
CC tumour cells that express KIM ligand can be inhibited with similarly
CC tagged KIM or anti-KIM ligand antibody
XX
SQ Sequence 334 AA;

Query Match 100.0%; Score 225; DB 2; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSYTTDGNVTSSDGLWNNQTLFLEHSLLTANTTKG 42
|||||
Db 249 PLSYTTDGNVTSSDGLWNNQTLFLEHSLLTANTTKG 290
|||||

RESULT 4
AAO26679
ID AAO26679 standard; protein; 334 AA.
XX
AC AAO26679;
XX
DT 20-MAR-2003 (first entry)
XX
DE KIM-1 related protein, SEQ ID NO 7.
XX
KW Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW renal disease; injury; renal cancer; human.
XX
OS Homo sapiens.
XX
XX WO200298920-A1.
XX
PD 12-DEC-2002.
XX
XX 31-MAY-2002; 2002WO-US017402.
XX
PF 01-JUN-2001; 2001US-0295449P.
PR 04-JUN-2001; 2001US-0295907P.
XX
XX (BIOJ) BIOGEN INC.
PA (GEO) GEN HOSPITAL CORP.
XX
PI Bailey V, Bonventre J;
XX
XX WPI; 2003-156845/15.
XX
PT New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
PS Disclosure; Page 39-40; 42pp; English.

CC The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a protein relating to the KIM-1
CC protein of the invention
XX
SQ Sequence 334 AA;

Query Match 100.0%; Score 225; DB 6; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;

Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLSYTTDGNVTSSDGLWNNQTLFLEHSLLTANTTKG 42
|||||
Db 249 PLSYTTDGNVTSSDGLWNNQTLFLEHSLLTANTTKG 290
|||||

RESULT 5
ADQ76696
ID ADQ76696 standard; protein; 334 AA.
XX
AC ADQ76696;
XX
DT 16-DEC-2004 (first entry)
XX
DE Human kidney injury molecule-1 (KIM-1).
XX
KW Human; KIM-1; kidney injury molecule-1; immunosuppressive; protozoacide;
KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW antiallergic; antiinflammatory; vasotropic; nephrotropic.
XX
OS Homo sapiens.
XX
XX
PH Key Location/Qualifiers
FT Peptide I..20
FT Modified-site /label= Signal_peptide
FT Modified-site 65..67
FT Modified-site /note= "N-glycosylated"
FT Modified-site 258..260
FT Modified-site /note= "N-glycosylated"
FT Modified-site 272..274
FT Modified-site /note= "N-glycosylated"
FT Modified-site 286..288
FT Modified-site /note= "N-glycosylated"
FT Domain 290..311
FT Region /note= "Transmembrane domain"
FT Region 324..334
FT /note= "C-terminal region unique to this splice variant"
XX
XX WO2004060041-A2.
XX
PD 22-JUL-2004.
XX
XX 29-DEC-2003; 2003WO-US041294.
XX
XX 30-DEC-2002; 2002US-0436934P.
XX
XX (BIOG-) BIOGEN IDEC MA INC.
XX
XX Rennert PD;
XX
XX WPI; 2004-534277/51.
XX
XX Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
XX Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
XX activation, or inhibiting production of a subset of antibodies against
XX one or more antigens.
XX
XX Disclosure; SEQ ID NO 1; 68pp; English.

CC The present sequence is that of human kidney injury molecule-1 (KIM-1), a
CC type I cell membrane glycoprotein and member of the T cell immunoglobulin
CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids or
CC 359 amino acids ADQ76690, depending on splice variation. It has been
CC discovered that treatment of a mammal with a KIM-1 antagonist alters the
CC interaction of T cells and other immune system cells, e.g. dendritic
CC cells, monocytes, macrophages and B cells, and thereby strongly
CC suppresses an IgG response to an antigen. Such treatment also eliminates
CC IgG1 production by memory B cells in response to subsequent challenge
CC with the antigen. Blockage of the binding of KIM-1 to its receptor
CC reduces secretion of interferon-gamma by immune cells engaged in an
CC antigen response in the mixed lymphocyte response assay. Based on these
CC discoveries, the invention provides methods for therapeutically

modulating immune function in autoimmune diseases and other disorders of the mammalian immune system. A method is claimed for inhibiting signalling between a T cell and a second cell, e.g. an antigen-presenting cell, in a mammal. The method involves identifying a mammal with an immune disease or disorder, or one preparing to receive a tissue graft, and administering a KIM-1 antagonist, especially a polypeptide comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a soluble polypeptide, which can include a KIM-1 mucin domain in addition to the KIM Ig domain, and may include a heterologous moiety such as an Fc moiety. The antagonist may be conjugated to a polymer such as polyethylene glycol. The KIM-1 antagonist is used in methods of: inhibiting activation of a B cell in a mammal; inhibiting disease relapse in an autoimmune disease; inhibiting epitope spreading in an autoimmune disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis, autoimmune haemolytic anaemia, Chagas disease, Grave's disease, idiopathic thrombocytopenia purpura, Wegener's granulomatosis, polyarteritis nodosa, rapidly progressive crescentic glomerulonephritis, graft-versus-host disease, or systemic lupus nephritis) or inflammatory disease or disorder (e.g. inflammatory bowel diseases such as Crohn's disease, ulcerative colitis, and ileitis); and inhibiting secretion of IFN-gamma by lymphocytes.

XX Sequence 334 AA;
Query Match 100.0%; Score 225; DB 8; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
|||||
DB 249 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 6
ADE36618
ID ADE36618 standard; protein; 339 AA.
XX
AC ADE36618;
XX
XX
DT 29-JAN-2004 (first entry)
XX
DE Plasmid pCR2.1-CG57008-03-S843 15B protein insert SEQ ID NO:28.
XX
KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.

OS Synthetic.
OS Homo sapiens.
XX WO2003080856-A2.
PN
PD 02-OCT-2003.
XX
PF 19-MAR-2003; 2003WO-US008490.
XX
PR 19-MAR-2002; 2002US-0365491P.
PR 13-SEP-2002; 2002US-0410618P.
XX
XX (CURA-) CURAGEN CORP.

XX
PI Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Meeri M;
PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;
XX
DR WPI; 2003-876927/81.
DR N-PSDB; ADE36617.
XX
XX New polypeptide, useful for preparing a composition for treating or preventing a pathology associated with NOVX polypeptide e.g. renal cancer or inflammation, or for tissue typing.
XX Example 1; SEQ ID NO 28; 239pp; English.

XX
CC The present invention describes an isolated human NOVX polypeptide, where X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at least 95 % identical to it, or a sequence comprising one or more conservative substitutions in the amino acid sequence. The NOVX polypeptide, and nucleic acid sequence encoding it, has cytostatic and antiinflammatory activities, and can be used in gene therapy, and in vaccines. The NOVX polypeptide is useful for preparing a composition for treating or preventing a pathology associated with NOVX polypeptide e.g. renal cancer or inflammation, or for tissue typing. The present sequence represents a pCR2.1-CG57008-03-S843 15B protein insert, which is used in an example from the present invention.

XX Sequence 339 AA;

Query Match 100.0%; Score 225; DB 7; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
|||||
DB 229 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 270

RESULT 7
ADE36592
ID ADE36592 standard; protein; 339 AA.
XX
AC ADE36592;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human NOV1a protein SEQ ID NO:2.
XX
KW human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.

OS Homo sapiens.
XX WO2003080856-A2.
PN
PD 02-OCT-2003.

XX
PF 19-MAR-2003; 2003WO-US008490.
XX
PR 19-MAR-2002; 2002US-0365491P.
PR 13-SEP-2002; 2002US-0410618P.

XX (CURA-) CURAGEN CORP.

XX
PI Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Meeri M;
PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;

XX
DR WPI; 2003-876927/81.
DR N-PSDB; ADE36591.

XX
PT New polypeptide, useful for preparing a composition for treating or preventing a pathology associated with NOVX polypeptide e.g. renal cancer or inflammation, or for tissue typing.

XX Claim 1; SEQ ID NO 2; 239pp; English.

XX
CC The present invention describes an isolated human NOVX polypeptide, where X is 1a to 1d or 2a to 2h, or its mature form, a sequence that is at least 95 % identical to it, or a sequence comprising one or more conservative substitutions in the amino acid sequence. The NOVX polypeptide, and nucleic acid sequence encoding it, has cytostatic and antiinflammatory activities, and can be used in gene therapy, and in vaccines. The NOVX polypeptide is useful for preparing a composition for treating or preventing a pathology associated with NOVX polypeptide e.g. renal cancer or inflammation, or for tissue typing. The present sequence represents human NOV1a from the present invention.

XX


```
SQ      Sequence 339 AA;
Query Match      100.0%; Score 225; DB 7; Length 339;
Best Local Similarity 100.0%; Pred. No. 1.9e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLYSYTTDGNVTWTESSDGLMNNQTLFLEHSLLTANTTKG 42
DB      229 PLYSYTTDGNVTWTESSDGLMNNQTLFLEHSLLTANTTKG 270

RESULT 8
ABR58582
ID      ABR58582 standard; protein; 359 AA.
XX
AC      ABR58582;
XX
DT      09-JUL-2003 (first entry)
XX
DE      Human cancer related protein SEQ ID NO:239.
XX
KW      Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW      heart disease; atherosclerosis; endometriosis.
XX
OS      Homo sapiens.
XX
PN      WO2003025138-A2.
XX
PD      27-MAR-2003.
XX
PF      17-SEP-2002; 2002WO-US029560.
XX
PR      17-SEP-2001; 2001US-0323469P.
PR      20-SEP-2001; 2001US-0323887P.
PR      13-NOV-2001; 2001US-0350666P.
PR      08-FEB-2002; 2002US-0355145P.
PR      08-FEB-2002; 2002US-0355257P.
PR      12-APR-2002; 2002US-0372246P.
XX
PA      (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI      Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI      Zlotnick A;
XX
WPI; 2003-354600/33.
N-PSDB; ACC72709.
XX
PT      New genes that are up-regulated or down-regulated in cancers, useful as
PT      markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT      therapeutic targets for screening drugs for treating these diseases.
XX
PS      Claim 12; Page 742; 767pp; English.
XX
CC      The present invention describes an isolated nucleic acid molecule, which
CC      comprises the sequence of any of the genes that are up-regulated or down-
CC      regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC      acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC      related gene nucleotide sequences which encode the proteins given in
CC      ABR58521 to ABR58709. Also described: (1) determining the presence or
CC      absence of a pathological cell in a patient; (2) an expression vector
CC      comprising a nucleic acid molecule described above; (3) a host cell
CC      comprising the vector; (4) an isolated polypeptide, which is encoded by
CC      the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC      of (4); (6) specifically targeting a compound to a pathological cell in a
CC      patient by administering to the patient the antibody above; and (7) a
CC      drug screening assay. The nucleic acid is useful as diagnostic markers or
CC      therapeutic targets. In particular, the nucleic acid is useful for
CC      diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC      bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC      pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC      atherosclerosis and endometriosis. The nucleic acid is also useful in
CC      drug screening, particularly for identifying agents for treating these
CC      pathologies
```

```
XX
SQ      Sequence 359 AA;
Query Match      100.0%; Score 225; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLYSYTTDGNVTWTESSDGLMNNQTLFLEHSLLTANTTKG 42
DB      249 PLYSYTTDGNVTWTESSDGLMNNQTLFLEHSLLTANTTKG 290

RESULT 9
ABP70439
ID      ABP70439 standard; protein; 359 AA.
XX
AC      ABP70439;
XX
DT      22-APR-2003 (first entry)
XX
DE      Amino acid sequence of human TIM-1 allele 1.
XX
KW      T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW      TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW      myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW      allergic T cell response; autoimmune disease.
XX
OS      Homo sapiens.
XX
PN      WO2003002722-A2.
XX
PD      09-JAN-2003.
XX
PF      01-JUL-2002; 2002WO-US020890.
XX
PR      29-JUN-2001; 2001US-0302344P.
XX
PA      (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI      McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
PI      WPI; 2003-210268/20.
DR      N-PSDB; ABZ68333.
XX
PT      New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT      Mucin domain gene sequences, useful for treating cancer or asthma,
PT      allergy, eczema or autoimmune disease.
XX
PS      Claim 10; Page 82; 94pp; English.
XX
CC      The present sequence is a human T cell immunoglobulin domain and mucin
CC      domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC      TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC      conserved IgV and mucin domains. The locus comprising the TIM family is
CC      genetically associated with immune dysfunction, including asthma. The TIM
CC      gene family is located within a region of human chromosome 5 that is
CC      commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC      of TIM-1 and TIM-3 are associated with susceptibility to airway
CC      hyperactivity and allergic T cell responses, and other variants
CC      associated with protection against these responses. T cells express TIM
CC      proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC      preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC      1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC      asthma, allergies, eczema or autoimmune diseases
XX
SQ      Sequence 359 AA;
Query Match      100.0%; Score 225; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PLYSYTTDGNVTWTESSDGLMNNQTLFLEHSLLTANTTKG 42
DB      249 PLYSYTTDGNVTWTESSDGLMNNQTLFLEHSLLTANTTKG 290
```

```
Db      249 PLYSYTGDNDVTVESSDGLMNNQTLFLEHSLLTANTTKG 290
RESULT 10
ID      ABP70438 standard; protein; 359 AA.
XX
XX
AC      ABP70438;
XX
DT      22-APR-2003 (first entry)
XX
DE      Amino acid sequence of human TIM-1 allele 1.
XX
KW      T cell immunoglobulin domain; mucin domain; TIM; TIM-1; TIM-2; TIM-3;
KW      TIM-4; immune dysfunction; chromosome 5; malignancy; allergy; eczema;
KW      myelodysplastic syndrome; airway hyperreactivity; cancer; asthma;
KW      allergic T cell response; autoimmune disease.
XX
OS      Homo sapiens.
XX
FN      WO2003002722-A2.
XX
PD      09-JAN-2003.
XX
PF      01-JUL-2002; 2002WO-US020890.
XX
PR      29-JUN-2001; 2001US-0302344P.
XX
PA      (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI      McIntire JJ, Dekruyff RH, Umetsu DT, Freeman GJ, Kuchroo V;
XX
DR      WPI; 2003-210268/20.
DR      N-PSDB; ABZ68332.
XX
XX      New nucleic acid comprising a mammalian T cell immunoglobulin domain and
PT      Mucin domain gene sequences, useful for treating cancer or asthma,
PT      allergy, eczema or autoimmune disease.
XX
PS      Claim 10; Page 80-81; 94pp; English.
XX
CC      The present sequence is a human T cell immunoglobulin domain and mucin
CC      domain-1 (TIM-1) polypeptide. The specification describes TIM-1, TIM-2,
CC      TIM-3 and TIM-4. TIM polypeptides are cell surface molecules with
CC      conserved IgV and mucin domains. The locus comprising the TIM family is
CC      genetically associated with immune dysfunction, including asthma. The TIM
CC      gene family is located within a region of human chromosome 5 that is
CC      commonly deleted in malignancies and myelodysplastic syndrome. Variants
CC      of TIM-1 and TIM-3 are associated with susceptibility to airway
CC      hyperactivity and allergic T cell responses, and other variants
CC      associated with protection against these responses. T cells express TIM
CC      proteins, which critically regulate CD4 T cell differentiation. Th1 cells
CC      preferentially express TIM-3, while Th2 cells preferentially express TIM-
CC      1. TIM polypeptides and polynucleotides are useful for treating cancer,
CC      asthma, allergies, eczema or autoimmune diseases
XX
SQ      Sequence 359 AA;
Query Match      100.0%; Score 225; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 PLYSYTGDNDVTVESSDGLMNNQTLFLEHSLLTANTTKG 42
DB      249 PLYSYTGDNDVTVESSDGLMNNQTLFLEHSLLTANTTKG 290
RESULT 11
ID      ABR48174 standard; protein; 359 AA.
XX
XX
AC      ABR48174;
XX
DT      20-MAR-2003 (first entry)
XX
DE      KIM-1 related protein, SEQ ID No 8.
XX
KW      Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW      proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW      renal disease; injury; renal cancer; human.
XX
Dt      12-JUN-2003 (first entry)
XX
DE      Human bladder cancer associated protein sequence SEQ ID NO:64.
XX
KW      Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS      Homo sapiens.
XX
FN      WO2003003906-A2.
XX
PD      16-JAN-2003.
XX
PF      03-JUL-2002; 2002WO-US021338.
XX
PR      03-JUL-2001; 2001US-0302814P.
PR      03-AUG-2001; 2001US-0310099P.
PR      08-NOV-2001; 2001US-0343705P.
PR      13-NOV-2001; 2001US-0350666P.
PR      12-APR-2002; 2002US-0372246P.
XX
XX      (EOSB-) EOS BIOTECHNOLOGY INC.
PA
XX
PI      Mack DH, Aziz N;
XX
DR      WPI; 2003-201532/19.
DR      N-PSDB; ACC50985.
XX
XX      Detecting a bladder cancer-associated transcript in a cell from a
PT      patient, comprises contacting a biological sample from the patient with a
PT      bladder cancer-associated polynucleotide or antibody.
XX
PS      Claim 10; Page 252; 307pp; English.
XX
CC      The present invention describes a method for detecting a bladder cancer-
CC      associated transcript in a cell from a patient. The method comprises
CC      contacting a biological sample from the patient with a polynucleotide
CC      that selectively hybridises to a sequence that is 80 % identical to a
CC      table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
CC      encode the human bladder cancer-associated proteins given in ABR48146 to
CC      ABR48242). Bladder cancer-associated sequences from the present invention
CC      have cytostatic activities, and can be used in antisense gene therapy and
CC      in vaccine production. The method can be used for detecting a bladder
CC      cancer-associated transcript in a cell from a patient. The method is
CC      useful in diagnosing or treating bladder cancer and in screening for
CC      compounds that modulate bladder cancer, such as hormones or antibodies.
CC      The nucleic acid molecules from the present invention may be used in
CC      various screening and diagnostic methods, and for gene therapy, vaccine
CC      and/or antisense/inhibition applications
XX
SQ      Sequence 359 AA;
Query Match      100.0%; Score 225; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      1 PLYSYTGDNDVTVESSDGLMNNQTLFLEHSLLTANTTKG 42
DB      249 PLYSYTGDNDVTVESSDGLMNNQTLFLEHSLLTANTTKG 290
RESULT 12
ID      AAO26680 standard; protein; 359 AA.
XX
XX
AC      AAO26680;
XX
DT      20-MAR-2003 (first entry)
XX
DE      KIM-1 related protein, SEQ ID No 8.
XX
KW      Cytostatic; gene therapy; antibody; antigen; antigen-binding;
KW      proteolytic release; inhibitor; kidney injury molecule-1; KIM-1;
KW      renal disease; injury; renal cancer; human.
XX
```

XX Homo sapiens.
OS WO200298920-A1.
PN
XX
XX 12-DEC-2002.
PD
XX 31-MAY-2002; 2002WO-US017402.
PF
XX 01-JUN-2001; 2001US-0295449P.
XX 04-JUN-2001; 2001US-0295907P.
PR
XX (BIOJ) BIOGEN INC..
PA (GEHO) GEN HOSPITAL CORP.
XX
XX Bailly V, Bonventre J;
XX WPI; 2003-156845/15.
DR
XX New antibody, antibody derivative or antigen-binding polypeptide that
PT inhibits proteolytic release of a soluble kidney injury molecule-1
PT polypeptide, useful for treating or preventing renal disease or injury,
PT e.g. renal cancer.
XX
XX Disclosure; Page 40-41; 42pp; English.
PS
XX The invention relates to a novel antibody, antibody derivative or antigen
CC -binding polypeptide that inhibits proteolytic release of a soluble
CC kidney injury molecule-1 (KIM-1) polypeptide from KIM-1-expressing cells.
CC The antibody, antibody derivative or antigen-binding polypeptide is
CC useful for treating or preventing renal disease or injury, e.g. renal
CC cancer. The antibody is also useful for inhibiting shedding of the KIM-1
CC polypeptide. This sequence represents a protein relating to the KIM-1
CC protein of the invention
XX
XX Sequence 359 AA;
SQ
Query Match 100.0%; Score 225; DB 6; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PLYSYTDCNDVTSSDGLWNNQQLFLEHSLLTANTTKG 42
DB 249 PLYSYTDCNDVTSSDGLWNNQQLFLEHSLLTANTTKG 290
RESULT 13
ADE36594
ID ADE36594 standard; protein; 359 AA.
XX
XX ADE36594;
AC
XX 29-JAN-2004 (first entry)
DT
XX Human NOV1b protein SEQ ID NO:4.
DE
XX human; NOVX; cytostatic; antiinflammatory; gene therapy; vaccine;
KW renal cancer; inflammation; tissue typing.
XX
XX Homo sapiens.
OS
XX WO2003080856-A2.
PN
XX 02-OCT-2003.
XX
XX 19-MAR-2003; 2003WO-US008490.
PF
XX 19-MAR-2002; 2002US-0365491P.
XX
XX 13-SEP-2002; 2002US-0410618P.
PR
XX (CURA-) CURAGEN CORP.
PA
XX Lepley DM, Rieger DK, Tse K, Rastelli L, Smithson G, Mesri M;
PI

PI Ooi CE, Anderson DW, Guo X, Giot L, Starling G;
XX WPI; 2003-876927/81.
DR N-PSDB; ADE36593.
XX
XX New polypeptide, useful for preparing a composition for treating or
PT preventing a pathology associated with NOVX polypeptide e.g. renal cancer
PT or inflammation, or for tissue typing.
PT
XX Claim 1; SEQ ID NO 4; 239pp; English.
PS
XX The present invention describes an isolated human NOVX polypeptide, where
CC X is la to ld or 2a to 2h, or its mature form, a sequence that is at
CC least 95 % identical to it, or a sequence comprising one or more
CC conservative substitutions in the amino acid sequence. The NOVX
CC polypeptide, and nucleic acid sequence encoding it, has cytostatic and
CC antiinflammatory activities, and can be used in gene therapy, and in
CC vaccines. The NOVX polypeptide is useful for preparing a composition for
CC treating or preventing a pathology associated with NOVX polypeptide e.g.
CC renal cancer or inflammation, or for tissue typing. The present sequence
CC represents human NOV1b from the present invention.
XX
XX Sequence 359 AA;
SQ
Query Match 100.0%; Score 225; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 PLYSYTDCNDVTSSDGLWNNQQLFLEHSLLTANTTKG 42
DB 249 PLYSYTDCNDVTSSDGLWNNQQLFLEHSLLTANTTKG 290
RESULT 14
ADN38984
ID ADN38984 standard; protein; 359 AA.
XX
XX ADN38984;
AC
XX 17-JUN-2004 (first entry)
DT
XX Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:302.
DE
XX Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulneryary; gene therapy; vaccine.
XX
XX Homo sapiens.
OS
XX WO2003042661-A2.
PN
XX 22-MAY-2003.
PD
XX 13-NOV-2002; 2002WO-US036810.
PF
XX 13-NOV-2001; 2001US-0350666P.
XX 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR

PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
DR N-PSDB; ADN38983.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX Claim 12; SEQ ID NO 302; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 225; DB 7; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTGDNDVTVESSDGLWNNNTQQLFLHSLTLTANTTKG 42
|||||
Db 249 PLYSYTGDNDVTVESSDGLWNNNTQQLFLHSLTLTANTTKG 290
|||||

RESULT 15
ADQ76690
ID ADQ76690 standard; protein; 359 AA.
XX
XX AC ADQ76690;
XX
XX 16-DEC-2004 (first entry)
DT
XX Human kidney injury molecule-1 (KIM-1).
DE
XX Human; KIM-1; kidney injury molecule-1; immunosuppressive; protozoacide;
KW muscular-gen.; neuroprotective; antianaemic; antithyroid; haemostatic;
KW antiallergic; antiinflammatory; vasotropic; nephrotropic.
XX
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH 1..20
FT Peptide /label= Signal_peptide
FT 65..67
FT Modified-site /note= "N-glycosylated"
FT 258..260
FT Modified-site /note= "N-glycosylated"
FT 272..274
FT Modified-site /note= "N-glycosylated"
FT

FT Modified-site 286..288
FT /note= "N-glycosylated"
FT Domain 290..311
FT /note= "Transmembrane domain"
FT Region 324..359
FT /note= "C-terminal region unique to this splice variant"
XX
XX WO2004060041-A2.
XX
XX 22-JUL-2004.
PD
XX 29-DEC-2003; 2003WO-US041294.
XX
XX 30-DEC-2002; 2002US-0436934P.
PR
XX (BIOG-) BIOGEN IDEC MA INC.
PA
XX Rennett PD;
PI
XX WPI; 2004-534277/51.
DR
XX Use of a kidney injury molecule-1 (KIM-1) antagonist for e.g. treating
XX Th2 cell-mediated diseases or inflammatory diseases, inhibiting B cell
PT activation, or inhibiting production of a subset of antibodies against
PT one or more antigens.
XX
XX Disclosure; SEQ ID NO 1; 68pp; English.
PS
XX The present sequence is that of human kidney injury molecule-1 (KIM-1), a
CC type I cell membrane glycoprotein and member of the T cell immunoglobulin
CC (Ig) and mucin domain (TIM) family. This protein has 334 amino acids
CC AD076696 or 359 amino acids, depending on splice variation. It has been
CC discovered that treatment of a mammal with a KIM-1 antagonist alters the
CC interaction of T cells and other immune system cells, e.g. dendritic
CC cells, monocytes, macrophages and B cells, and thereby strongly
CC suppresses an IgG response to an antigen. Such treatment also eliminates
CC IgG1 production by memory B cells in response to subsequent challenge
CC with the antigen. Blockage of the binding of KIM-1 to its receptor
CC reduces secretion of interferon-gamma by immune cells engaged in an
CC antigen response in the mixed lymphocyte response assay. Based on these
CC discoveries, the invention provides methods for therapeutically
CC modulating immune function in autoimmune diseases and other disorders of
CC the mammalian immune system. A method is claimed for inhibiting
CC signalling between a T cell and a second cell, e.g. an antigen-presenting
CC cell, in a mammal. The method involves identifying a mammal with an
CC immune disease or disorder, or one preparing to receive a tissue graft,
CC and administering a KIM-1 antagonist, especially a polypeptide
CC comprising: a KIM-1 Ig domain, and lacking a transmembrane domain and a
CC KIM-1 cytoplasmic domain; an anti-KIM-1 antibody; or an antigen-binding
CC fragment of an anti-KIM-1 antibody. Preferably, the KIM-1 antagonist is a
CC soluble polypeptide, which can include a KIM-1 mucin domain in addition
CC to the KIM Ig domain, and may include a heterologous moiety such as an Fc
CC moiety. The antagonist may be conjugated to a polymer such as
CC polyethylene glycol. The KIM-1 antagonist is used in methods of:
CC inhibiting activation of a B cell in a mammal; inhibiting in an autoimmune
CC in an autoimmune disease; inhibiting epitope spreading in an autoimmune
CC disease; treating a Th2 cell-mediated disease (e.g. myasthenia gravis,
CC autoimmune haemolytic anaemia, Chagas disease, Grave's disease,
CC idiopathic thrombocytopenia purpura, Wegener's granulomatosis,
CC polyarthritis nodosa, rapidly progressive crescentic glomerulonephritis,
CC graft-versus-host disease, or systemic lupus nephritis) or inflammatory
CC disease or disorder (e.g. inflammatory bowel diseases such as Crohn's
CC disease, ulcerative colitis, and ileitis); and inhibiting secretion of
CC IFN-gamma by lymphocytes.
XX
SQ Sequence 359 AA;

Query Match 100.0%; Score 225; DB 8; Length 359;
Best Local Similarity 100.0%; Pred. No. 2e-20;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTGDNDVTVESSDGLWNNNTQQLFLHSLTLTANTTKG 42
|||||

Db 249 PLYSYTTDGN DTVTESSDGLWNNNQTLFLEHSLLTANTTKG 290

Search completed: June 29, 2005, 09:09:07
Job time : 30.8703 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 5.77603 Seconds
(without alignments)
542.805 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225

Sequence: 1 PLYSYTGDNDVTBSSDGL.....NNQTQLFLHSLLTANTTKG 42

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCUTUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	100.0	334	4	US-09-197-970B-7
2	178	79.1	451	1	US-08-287-001A-2
3	178	79.1	451	5	PCR-US95-09941-2
4	60	26.7	602	1	US-08-168-091A-2
5	60	26.7	602	1	US-08-428-926-5
6	60	26.7	602	1	US-08-428-927-5
7	60	26.7	602	1	US-08-428-298-5
8	60	26.7	602	1	US-08-339-517-5
9	60	26.7	606	4	US-09-603-208A-56
10	60	26.7	1070	3	US-08-697-954-2
11	58	25.8	981	4	US-09-991-258-13
12	57	25.3	307	4	US-09-197-970B-3
13	54.5	24.2	323	6	5260223-4
14	54.5	24.2	323	6	5260223-4
15	53.5	23.8	10182	3	US-09-134-001C-3159
16	53	23.6	190	3	US-08-341-018-44
17	53	23.6	190	3	US-08-470-335-217
18	53	23.6	190	3	US-08-470-339-217
19	53	23.6	190	4	US-08-467-602-411
20	53	23.6	210	4	US-08-467-602-191
21	53	23.6	210	4	US-08-411-295F-315
22	53	23.6	213	4	US-08-467-603-186
23	53	23.6	213	4	US-08-411-295F-310
24	53	23.6	219	4	US-08-467-602-200
25	53	23.6	219	4	US-08-411-295F-324
26	53	23.6	222	4	US-08-467-602-197
27	53	23.6	222	4	US-08-411-295F-321

28	53	23.6	224	4	US-09-248-796A-18758	Sequence 18758, A
29	53	23.6	233	4	US-08-467-602-194	Sequence 194, App
30	53	23.6	233	4	US-08-411-295F-318	Sequence 318, App
31	53	23.6	242	4	US-08-467-602-203	Sequence 203, App
32	53	23.6	242	4	US-08-411-295F-327	Sequence 327, App
33	53	23.6	309	4	US-09-248-796A-25339	Sequence 25339, A
34	53	23.6	352	4	US-08-467-602-239	Sequence 239, App
35	53	23.6	352	4	US-08-411-295F-165	Sequence 165, App
36	53	23.6	355	4	US-08-467-602-234	Sequence 234, App
37	53	23.6	355	4	US-08-411-295F-160	Sequence 160, App
38	53	23.6	361	4	US-08-467-602-248	Sequence 248, App
39	53	23.6	361	4	US-08-411-295F-174	Sequence 174, App
40	53	23.6	364	4	US-08-467-602-245	Sequence 245, App
41	53	23.6	364	4	US-08-411-295F-171	Sequence 171, App
42	53	23.6	375	4	US-08-467-602-242	Sequence 242, App
43	53	23.6	375	4	US-08-411-295F-168	Sequence 168, App
44	53	23.6	384	4	US-08-467-602-251	Sequence 251, App
45	53	23.6	384	4	US-08-411-295F-177	Sequence 177, App

ALIGNMENTS

RESULT 1
US-09-197-970B-7
; Sequence 7, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-No. 6664385-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-197-970B-7
Query Match 100.0%; Score 225; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 2.8e-22;

```
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTGDNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
    |||||
Db 249 PLYSYTGDNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 2
US-08-287-001A-2
; Sequence 2, Application US/08287001A
; Patent No. 5622861
; GENERAL INFORMATION:
; APPLICANT: KAPLAN, GERARDO
; APPLICANT: FRINSTONE, STEPHEN M.
; TITLE OF INVENTION: HEPATITIS A VIRUS RECEPTOR AND METHODS
; TITLE OF INVENTION: OF USE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NEEDLE & ROSENBERG, P.C.
; STREET: Suite 1200, The Candler Bldg, 127 Peachtree
; STREET: Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/287,001A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-9880
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09941-2

Query Match 79.1%; Score 178; DB 1; Length 451;
Best Local Similarity 81.0%; Pred. No. 8.8e-16;
Matches 34; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTGDNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
    |||||
Db 329 PLYSYTGDNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 370

RESULT 3
US-08-287-001A-2
; Sequence 2, Application PC/TUS9509941
; Patent No. 5665862
; GENERAL INFORMATION:
; APPLICANT: FISCHBACH, GERALD.
; APPLICANT: FALLS, DOUGLAS R.
; APPLICANT: ROSEN, KENNETH M.
; APPLICANT: CORFAS, GABRIEL
; TITLE OF INVENTION: Neurotrophic Factor
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE AND COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/168,091A
; FILING DATE: 15-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/953,742
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: HMI-002CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
```

```
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/09941
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/287,001
; FILING DATE: 5 AUG 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Spratt, Gwedolyn D.
; REGISTRATION NUMBER: 36,016
; REFERENCE/DOCKET NUMBER: 1414.621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404/688-0770
; TELEFAX: 404/688-9880
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 451 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-09941-2

Query Match 79.1%; Score 178; DB 5; Length 451;
Best Local Similarity 81.0%; Pred. No. 8.8e-16;
Matches 34; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PLYSYTGDNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
    |||||
Db 329 PLYSYTGDNDVTSSDGLWNNNTQLFLEHSLLTANTTKG 370

RESULT 4
US-08-168-091A-2
; Sequence 2, Application US/08168091A
; Patent No. 5665862
; GENERAL INFORMATION:
; APPLICANT: FISCHBACH, GERALD.
; APPLICANT: FALLS, DOUGLAS R.
; APPLICANT: ROSEN, KENNETH M.
; APPLICANT: CORFAS, GABRIEL
; TITLE OF INVENTION: Neurotrophic Factor
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE AND COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/168,091A
; FILING DATE: 15-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/953,742
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: HMI-002CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
```



```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/339517
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 853D2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-428-298-5

Query Match 26.7%; Score 60; DB 1; Length 602;
Best Local Similarity 31.7%; Pred. No. 10;
Matches 13; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 3 YSYTGDNDVTYESSDGLWNNNQTLFL--EHSLLTANTTK 41
Db 301 YTSHTHSMVTQTTPSHWSNGHTESILSESHSVLVSSSVE 341

RESULT 8
US-08-339-517-5
; Sequence 5, Application US/08339517
; Patent No. 5770567
; GENERAL INFORMATION:
; APPLICANT: Ho, Wei-Hsien
; TITLE OF INVENTION: SENSORY AND MOTOR NEURON DERIVED FACTOR (SMDF)
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/339,517
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 853
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 602 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-339-517-5

Query Match 26.7%; Score 60; DB 1; Length 602;
Best Local Similarity 31.7%; Pred. No. 10;
```

```
Matches 13; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 3 YSYTGDNDVTYESSDGLWNNNQTLFL--EHSLLTANTTK 41
Db 301 YTSHTHSMVTQTTPSHWSNGHTESILSESHSVLVSSSVE 341

RESULT 9
US-09-603-208A-56
; Sequence 56, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; TITLE OF INVENTION: RESISTANCE AND TOLERANCE PROTEINS
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 56
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-56

Query Match 26.7%; Score 60; DB 4; Length 606;
Best Local Similarity 34.1%; Pred. No. 10;
Matches 14; Conservative 4; Mismatches 23; Indels 0; Gaps 0;

QY 1 PLYSYTGDNDVTYESSDGLWNNNQTLFLSHSLTANTTK 41
Db 559 FLVEYQKEGDMFMGMDKIKEETVQLFLSASSSSSKTRK 599

RESULT 10
US-08-697-954-2
; Sequence 2, Application US/08697954
; Patent No. 6284535
; GENERAL INFORMATION:
; APPLICANT: Role, Lorna W.
; TITLE OF INVENTION: SPLICE VARIANTS OF THE HEREGULIN GENE, PARTA, AND
; TITLE OF INVENTION: US THEREOF
; NUMBER OF SEQUENCES: 4
```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/697,954
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 46839-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1070 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-697-954-2

```

Query Match      26.7%; Score 60; DB 3; Length 1070;
Best Local Similarity 31.7%; Pred.No. 20;
Matches 13; Conservative 11; Mismatches 15; Indels 2; Gaps 1;

QY 3 YSYTTDNDVTVSSDGLWNNNQLEF--EHSLLTANTTK 41
   |||:::|||||:::|
Db 586 YTGTHHSMTVOTPPSHSWNGHTESILSHSHVLSVSSVE 626

```

```

RESULT 11
US-09-991-258-13
; Sequence 13, Application US/09991258
; Patent No. 6783939
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanson, Ronald
; TITLE OF INVENTION: ALPHAVIRUS
; TITLE OF INVENTION: VACCINES
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 981
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; No. 6783939e =
US-09-991-258-13

```

```

Query Match      25.8%; Score 58; DB 4; Length 981;
Best Local Similarity 47.6%; Pred. NO. 34;
Matches 10; Conservative 4; Mismatches 7; Indels

QY      11 DTVTSSDGLWNNNQTLFLFLE 31
      : ||| ||||| :
DB      484 ETTWESLDHLWNNNQMFWIQ 504
      : ||| ||||| :

RESULT 12
US-09-197-970B-3
; Sequence 3, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
;

```

THURGOOD J. CAULFIELD
 TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Biogen, Inc.
 STREET: 14 Cambridge Center
 CITY: Cambridge
 STATE: MA
 COUNTRY: USA
 ZIP: 02142

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,970B
FILING DATE: 23-NO. 6664385-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,228

FILING DATE: 24-MAY-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Levine, Leslie M.
 REGISTRATION NUMBER: 35,245
 REFERENCE/DOCKET NUMBER: A010 PCT CIP
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 679-2810
 TELEFAX: (617) 679-2838
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 307 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:
 7-970B-3

Query Match 25.3%; Score 57; DB 4; Length 307;
Best Local Similarity 45.0%; Pred.No. 11;
Matches 18; Conservative 4; Mismatches 14; Indels

QY 4 SYT-TDGN¹DTV²ESSDGLWNNQ³QLFLEHSLLTANTTKG 42

Db 199 SYTPADWNG¹VT--SSEAWNNH²TVRIPLRKP--QRNP³TKG 235

RESULT 13
5260223-4
; Patent No. 5260223
; APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN,
JOHN G.; IP, STEPHEN H.; KRANGEL, MICHAEL S.
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA
T CELL RECEPTOR

```
;
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/187,698
; FILING DATE: 29-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,256
; FILING DATE: 29-OCT-1987
; APPLICATION NUMBER: 16,252
; FILING DATE: 19-FEB-1987
; APPLICATION NUMBER: 882,100
; FILING DATE: 03-JUL-1986
; SEQ ID NO:4:
; LENGTH: 323
5260223-4

Query Match      24.2%; Score 54.5; DB 6; Length 323;
Best Local Similarity 30.0%; Pred. No. 26;
Matches 12; Conservative 6; Mismatches 19; Indels 3; Gaps 1;

QY      1 PLYSYTDGNDVTVESSDGLWNNQTLFLEHSLLTANTT 40
Db      251 PKYNSKDANDVITWDPKDNWSKDANDTLL---LQLTNTS 287

RESULT 14
5260223-4
; Patent No. 5260223
; APPLICANT: BRENNER, MICHAEL B.; STROMINGER, JACK L.; SEIDMAN,
; JOHN G.; IP, STEPHEN H.; KRANGEL, MICHAEL S.
; TITLE OF INVENTION: METHODS FOR DETECTION OF HUMAN GAMMA
; T CELL RECEPTOR
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/187,698
; FILING DATE: 29-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,256
; FILING DATE: 29-OCT-1987
; APPLICATION NUMBER: 16,252
; FILING DATE: 19-FEB-1987
; APPLICATION NUMBER: 882,100
; FILING DATE: 03-JUL-1986
; SEQ ID NO:4:
; LENGTH: 323
5260223-4

Query Match      24.2%; Score 54.5; DB 6; Length 323;
Best Local Similarity 30.0%; Pred. No. 26;
Matches 12; Conservative 6; Mismatches 19; Indels 3; Gaps 1;

QY      1 PLYSYTDGNDVTVESSDGLWNNQTLFLEHSLLTANTT 40
Db      251 PKYNSKDANDVITWDPKDNWSKDANDTLL---LQLTNTS 287

RESULT 15
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
```

```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match      23.8%; Score 53.5; DB 3; Length 10182;
Best Local Similarity 26.0%; Pred. No. 2.5e+03;
Matches 13; Conservative 13; Mismatches 13; Indels 11; Gaps 2;

QY      3 YSYTTDGDND-----TVTESSDGLWNNQTLFLEHS-----LLTANTTK 41
Db      2418 YTEKLNGETOKSFITTKNNNGKWTINKPKYVEFNQDNGKVVSANTIK 2467
```

Search completed: June 29, 2005, 09:02:57
Job time : 6.77603 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:58:20 ; Search time 26.3222 Seconds
(without alignments)
613.589 Million cell updates/sec

Title: US-10-718-321-6_COPY_40_81

Perfect score: 225
Sequence: 1 PLYSVTTGNDVTWSSDGL.....NNQTQLEHSLLTANTTKG 42

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1717557 seqs, 384547976 residues

Total number of hits satisfying chosen parameters: 1717557

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications_AA.*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	100.0	81	17	US-10-718-321-6
2	225	100.0	334	17	US-10-655-506-7
3	225	100.0	334	17	US-10-718-321-7
4	225	100.0	339	17	US-10-391-939A-2
5	225	100.0	339	17	US-10-391-939A-28
6	225	100.0	359	14	US-10-188-012-17
7	225	100.0	359	14	US-10-188-012-17
8	225	100.0	359	15	US-10-188-012-19
9	225	100.0	359	15	US-10-295-027-302
10	225	100.0	359	17	US-10-188-832-64
11	225	100.0	359	17	US-10-391-939A-4
					US-10-663-497-17

12	225	100.0	359	17	US-10-663-497-19
13	225	100.0	359	17	US-10-718-321-8
14	225	100.0	359	17	US-10-847-918-25
15	225	100.0	364	14	US-10-188-012-25
16	225	100.0	364	17	US-10-663-497-25
17	225	100.0	365	14	US-10-188-012-21
18	225	100.0	365	17	US-10-663-497-21
19	220	97.8	359	14	US-10-188-012-23
20	220	97.8	359	17	US-10-663-497-23
21	220	97.8	364	14	US-10-188-012-27
22	220	97.8	364	17	US-10-663-497-27
23	189	84.0	263	17	US-10-391-939A-6
24	189	84.0	263	17	US-10-391-939A-32
25	189	84.0	263	17	US-10-391-939A-39
26	189	84.0	263	17	US-10-805-177-50
27	99	44.0	18	17	US-10-718-321-1
28	58	25.8	272	16	US-10-425-115-288232
29	58	25.8	279	16	US-10-425-115-343835
30	58	25.8	981	9	US-09-991-258-13
31	58	25.8	981	18	US-10-929-234-13
32	57	25.3	307	17	US-10-655-506-3
33	57	25.3	658	15	US-10-335-977-9518
34	57	25.3	679	15	US-10-369-493-20059
35	57	25.3	996	15	US-10-335-977-9519
36	57	25.3	2399	15	US-10-282-122A-59054
37	56.5	25.1	666	16	US-10-437-963-196644
38	56.5	25.1	727	16	US-10-296-723A-14
39	56	24.9	145	16	US-10-437-963-127039
40	56	24.9	338	15	US-10-425-114-37036
41	56	24.9	338	15	US-10-412-698B-514
42	56	24.9	554	14	US-10-125-692-21
43	56	24.9	584	15	US-10-369-493-19944
44	55.5	24.7	125	15	US-10-424-599-220078
45	55	24.4	379	16	US-10-408-765A-1975

ALIGNMENTS

RESULT 1
US-10-718-321-6
; Sequence 6, Application US/10718321
; Publication NO. US20050112117A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; FILE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; PRIOR FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-6

Query Match 100.0%; Score 225; DB 17; Length 81;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PLYSVTTGNDVTWSSDGLWNNQTQLEHSLLTANTTKG 42
|||||

```
Db      40  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 81

RESULT 2
US-10-655-506-7
; Sequence 7, Application US/10655506
; Publication No. US20050089868A1
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
;           Joseph V. Bonventre
;           Catherine A. Hession
;           Takaharu Ichimura
;           Henry Wei
;           Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/655,506
; FILING DATE: 04-Sep-2003
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970
; FILING DATE: 23-Nov-1998
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-10-655-506-7

Query Match      100.0%; Score 225; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 5.6e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 42
Db      249  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 290

RESULT 3
US-10-718-321-7
; Sequence 7, Application US/10718321
; Publication No. US2005011217A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; TITLE OF INVENTION: Shedding of KIM-1
```

```
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 334
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-7

Query Match      100.0%; Score 225; DB 17; Length 334;
Best Local Similarity 100.0%; Pred. No. 5.6e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 42
Db      249  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 290

RESULT 4
US-10-391-939A-2
; Sequence 2, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Giot, Loic;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Mesri, Mehdi;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glennda;
; APPLICANT: Starling, Gary;
; APPLICANT: Tse, Kam-Fai
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-048
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 2
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-2

Query Match      100.0%; Score 225; DB 17; Length 339;
Best Local Similarity 100.0%; Pred. No. 5.7e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 42
Db      229  PLYSYTTDGDNDVTSSDGLWNNNQTLFLEHSLLTANTTKG 270

RESULT 5
US-10-391-939A-28
; Sequence 28, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
```

```

; APPLICANT: Giot, Loic ;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
; APPLICANT: Mesri, Mehdi ;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca ;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glennda ;
; APPLICANT: Starling, Gary ;
; APPLICANT: Tse, Kam-Fai
; FILE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 28
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-28

Query Match      100.0%; Score 225; DB 17; Length 339;
Best Local Similarity 100.0%; Pred. No. 5.7e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
Db 229 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 270

RESULT 6
US-10-188-012-17
; Sequence 17, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE OF INVENTION: Use Thereof
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-188-012-17

Query Match      100.0%; Score 225; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
Db 249 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 7
US-10-188-012-19
; Sequence 19, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE OF INVENTION: Use Thereof
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-188-012-19

Query Match      100.0%; Score 225; DB 14; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 42
Db 249 PLYSYTTDGNVTVTSSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 8
US-10-295-027-302
; Sequence 302, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/563,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
```

```
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 302
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-302

Query Match          100.0%; Score 225; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 42
|||||
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 9
US-10-188-832-64
; Sequence 64, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-0023300S
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-64

Query Match          100.0%; Score 225; DB 15; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 42
|||||
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 10
US-10-391-939A-4
; Sequence 4, Application US/10391939A
; Publication No. US20050014687A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Giot, Loic ;
; APPLICANT: Guo, Xiaojia Sasha;
; APPLICANT: Lepley, Denise M.;
```

```
; APPLICANT: Mesri, Mehdi ;
; APPLICANT: Ooi, Chean Eng;
; APPLICANT: Rastelli, Luca ;
; APPLICANT: Rieger, Daniel K.;
; APPLICANT: Smithson, Glenda ;
; APPLICANT: Starling, Gary ;
; APPLICANT: Tse, Kam-Fai
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
; FILE REFERENCE: 21402-048
; CURRENT APPLICATION NUMBER: US/10/391,939A
; CURRENT FILING DATE: 2003-03-19
; PRIOR APPLICATION NUMBER: 60/365,491
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 60/410,618
; PRIOR FILING DATE: 2002-09-13
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 4
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-391-939A-4

Query Match          100.0%; Score 225; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 42
|||||
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 11
US-10-663-497-17
; Sequence 17, Application US/10663497
; Publication No. US20050095593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; TITLE OF INVENTION: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT APPLICATION NUMBER: US/10/663,497
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 10/188,012
; PRIOR FILING DATE: 2002-07-01
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(360)
; OTHER INFORMATION: TIM-1 allele 1
US-10-663-497-17

Query Match          100.0%; Score 225; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 42
|||||
DB 249 PLYSYTTDGDNDVTVESSDGLWNNNTQLFLEHSLLTANTTKG 290

RESULT 12
US-10-663-497-19
```



```
; Sequence 19, Application US/10663497
; Publication No. US2005009593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; TITLE OF INVENTION: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/663,497
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 10/188,012
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-663-497-19

Query Match 100.0%; Score 225; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 42
DB 249 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 290

RESULT 13
US-10-718-321-8
; Sequence 8, Application US/10718321
; Publication No. US2005011217A1
; GENERAL INFORMATION:
; APPLICANT: Biogen Idec Ma Inc.
; APPLICANT: Bailly, Veronique
; APPLICANT: Bonventre, Joseph
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: Molecules and Methods for Inhibiting
; TITLE OF INVENTION: Shedding of KIM-1
; FILE REFERENCE: A124 US
; CURRENT APPLICATION NUMBER: US/10/718,321
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: 60/295449
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: 60/295907
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: PCT/US02/17402
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-718-321-8

Query Match 100.0%; Score 225; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 42
DB 249 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 290

; Sequence 19, Application US/10663497
; Publication No. US2005009593A1
; GENERAL INFORMATION:
; APPLICANT: Dale Umetsu
; APPLICANT: Rosemarie DeKruyff
; APPLICANT: Jennifer McIntire
; APPLICANT: Gordon Freeman
; TITLE OF INVENTION: T CELL REGULATORY GENES ASSOCIATED WITH
; TITLE OF INVENTION: IMMUNE DISEASE
; FILE REFERENCE: STAN-235CIP
; CURRENT FILING DATE: 2003-09-15
; PRIOR APPLICATION NUMBER: 60/663,497
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 10/188,012
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 359
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(359)
; OTHER INFORMATION: TIM-1, allele 2
US-10-663-497-19

Query Match 100.0%; Score 225; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 42
DB 249 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 290

RESULT 14
US-10-847-918-25
; Sequence 25, Application US/10847918
; Publication No. US20050119210A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Be, Xiaobing
; APPLICANT: Liu, Wei
; APPLICANT: Slonim, Donna
; APPLICANT: Howes, Steve
; TITLE OF INVENTION: Compositions and Methods for Diagnosing and Treating Cancers
; FILE REFERENCE: 031896-026000 (AM101264)
; CURRENT APPLICATION NUMBER: US/10/847,918
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US 60/471,729
; PRIOR FILING DATE: 2003-05-20
; NUMBER OF SEQ ID NOS: 14937
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-847-918-25

Query Match 100.0%; Score 225; DB 17; Length 359;
Best Local Similarity 100.0%; Pred. No. 6.2e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 42
DB 249 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 290

RESULT 15
US-10-188-012-25
; Sequence 25, Application US/10188012
; Publication No. US20030124114A1
; GENERAL INFORMATION:
; APPLICANT: McIntire, Jennifer Jones
; APPLICANT: Umetsu, Dale T.
; APPLICANT: Dekruyff, Rosemarie
; APPLICANT: Kuchroo, Vijay
; APPLICANT: Freeman, Gordon J.
; TITLE OF INVENTION: T Cell Regulatory Genes and Methods of
; FILE REFERENCE: STAN-235
; CURRENT APPLICATION NUMBER: US/10/188,012
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: 60/302,344
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 364
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(364)
; OTHER INFORMATION: TIM-1 allele 5
US-10-188-012-25

Query Match 100.0%; Score 225; DB 14; Length 364;
Best Local Similarity 100.0%; Pred. No. 6.3e-22;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 42
DB 254 PLYSYTTCGNDVTSSDGLWNNQTLFLEHSLLTANTTKG 295

Search completed: June 29, 2005, 09:33:00
```

Job time : 26.3222 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 29, 2005, 08:57:16 ; Search time 32.4637 Seconds
(without alignments)
1277.688 Million cell updates/sec

Title: US-10-718-321-6
Perfect score: 432
Sequence: 1 MELPRQNHPEVATSSPQP.....NNQTQLFLEHSLLTANTTKG 81

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03: *
1: uniprot_sprot: *
2: uniprot_trembl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	432	100.0	359	2	Q43656
2	432	100.0	364	2	Q96D42
3	348	80.6	451	2	Q95144
4	348	80.6	460	2	O18984
5	335	77.5	469	2	Q7J348
6	335	77.5	473	2	Q7J347
7	335	77.5	474	2	Q46597
8	335	77.5	478	2	O46598
9	89	20.6	307	2	O54947
10	86	20.4	153	2	Q7L9D0
11	86	19.9	3150	2	Q7PMD5
12	83	19.2	542	2	Q75UG6
13	82.5	19.1	684	2	Q88SZ4
14	79.5	18.4	965	2	Q22286
15	79	18.3	681	2	Q9PFW5
16	77	17.8	440	2	Q94J16
17	77	17.8	1139	2	Q6NTK5
18	76.5	17.7	282	2	Q8VIM1
19	76.5	17.7	282	2	Q7TPU2
20	76	17.6	875	2	Q9H706
21	76	17.6	1006	2	Q62901
22	75.5	17.5	224	2	Q9CET7
23	75.5	17.5	303	2	Q9D116
24	75.5	17.5	303	2	Q8R3W0
25	75.5	17.5	319	2	Q8BP27
26	75.5	17.5	319	2	Q9D0D7
27	75.5	17.5	319	2	Q9D4W4
28	75.5	17.5	1146	2	Q8X0L3
29	75	17.4	670	2	Q6BVR2
30	75	17.4	777	2	Q9BSF5
31	75	17.4	877	2	Q8ND03

32	74.5	17.2	298	2	Q9U5E7
33	74.5	17.2	1015	2	Q86AG0
34	73.5	17.0	823	2	Q75IH8
35	73.5	17.0	2819	2	Q8BEN8
36	73	16.9	579	2	Q69P50
37	73	16.9	709	2	Q86A80
38	73	16.9	982	2	Q6CSV5
39	73	16.9	1423	2	Q7YS39
40	72.5	16.8	338	2	Q9SQQ6
41	72.5	16.8	634	2	Q7RJ33
42	72.5	16.8	760	1	ABB2 MOUSE
43	72.5	16.8	837	2	Q7QEG1
44	72.5	16.8	888	2	Q6FB73
45	72.5	16.8	1082	2	Q6PDP0

ALIGNMENTS

RESULT 1
O43656 PRELIMINARY; PRT; 359 AA.
ID O43656
AC O43656;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT - 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hepatitis A virus cellular receptor 1.
GN Name=HAVCr-1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98325180; PubMed=9658108;
RA Feigelsstock D., Thompson P., Mattoo P., Zhang Y., Kaplan G.G.;
RT "The human homolog of HAVcr-1 codes for a hepatitis A virus cellular receptor.";
RL J. Virol. 72:6621-6628(1998).
DR EMBL; AF043724; AAC39862.1; -;
DR GO; GO:0004872; Fireceptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG_1.
DR PROSITE; PS00835; IG_LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 359 AA; 38704 MW; C207FEC562DC62CA CRC64;

Query Match 100.0%; Score 432; DB 2; Length 359;
Best Local Similarity 100.0%; Pred.No. 2.9e-33;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MELPRQNHPEVATSSPQP	ATHPTTLQCAIRREPTSSPLSYTTDGN	DTVTESDGLW 60
Db	210	MELPRQNHPEVATSSPQP	ATHPTTLQCAIRREPTSSPLSYTTDGN	DTVTESDGLW 269

QY	61	NNQTQLFLEHSLLTANTTKG	81
Db	270	NNQTQLFLEHSLLTANTTKG	290

RESULT 2
Q96D42 PRELIMINARY; PRT; 364 AA.
ID Q96D42
AC Q96D42;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE HAVCr1 protein.
GN Name=HAVCr1;

```
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Peltingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Schenker C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J.M., Hsieh F.,
RA Diatchenko L., Marudina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL: BC013325; AAH13325.1; -
DR EMBL: CR457114; CAG33395.1; -
DR GenBank: HGNC:17866; HAVCR1.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003006; Ig_MHC.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 364 AA; 39249 MW; AA67C7DC7FAC81F1 CRC64;

Query Match 100.0%; Score 432; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 3e-33;
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPLPRQNHFPVATSPSSPPQAEHTPTTLOGAIRREPTSSPLSYTTGNDVTVESDGLW 60
Db 215 MPLPRQNHFPVATSPSSPPQAEHTPTTLOGAIRREPTSSPLSYTTGNDVTVESDGLW 274

QY 61 NNQTLFLEHSLLTANTTKG 81
Db 275 NNQTLFLEHSLLTANTTKG 295

RESULT 3
QY5144 ID Q95144 PRELIMINARY; PRT; 451 AA.
AC Q95144;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE HAVcr-1 protein precursor.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
```

```
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97015129; PubMed=8861957;
RA Kaplan G., Totusuka A., Thompson P., Akatsuka T., Moritsugu Y.,
RA Feinstein S.M.;
RT "Identification of a surface glycoprotein on African green monkey
RT kidney cells as a receptor for hepatitis A virus.";
RL EMBL: X98252; CAA66906.1; -
DR PIR: S71754; S71754.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL. 1 17 Potential.
SQ SEQUENCE 451 AA; 48774 MW; 5D395C5455AA4332 CRC64;

Query Match 80.6%; Score 348; DB 2; Length 451;
Best Local Similarity 82.5%; Pred. No. 4.3e-25;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 PLPRQNHFPVATSPSSPPQAEHTPTTLOGAIRREPTSSPLSYTTGNDVTVESDGLW 61
Db 291 PLPRQNHFPVATSPSSPPQAEHTPTTLOGAIRREPTSSPLSYTTGNDVTVESDGLW 350

QY 62 NNQTLFLEHSLLTANTTKG 81
Db 351 NNQTLFLEHSLLTANTTKG 370

RESULT 4
QY5144 ID Q18984 PRELIMINARY; PRT; 460 AA.
AC Q18984;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Hepatitis A virus receptor.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97368416; PubMed=9225030;
RA Ashida M., Hamada C.;
RT "Molecular cloning of the hepatitis A virus receptor from a simian
RT cell line.";
RL J. Gen. Virol. 78:1565-1569(1997).
DR EMBL: D88585; BAA21556.1; -
DR GO: GO:0004872; F:receptor activity; IEA.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR007110; Ig-like.
DR SMART: SM00409; IG; 1.
DR PROSITE: PS50835; IG LIKE; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 460 AA; 49687 MW; 1D1A0C1D832139EA CRC64;

Query Match 80.6%; Score 348; DB 2; Length 460;
Best Local Similarity 82.5%; Pred. No. 4.4e-25;
Matches 66; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 2 PLPRQNHFPVATSPSSPPQAEHTPTTLOGAIRREPTSSPLSYTTGNDVTVESDGLW 61
Db 300 PLPRQNHFPVATSPSSPPQAEHTPTTLOGAIRREPTSSPLSYTTGNDVTVESDGLW 359
```

```

QY 62 NNQTLFLEHSLLTANTTKG 81
DB 360 NNQTLSPHSPQMVNTTEG 379

RESULT 5
Q7JJ48
ID AC Q7JJ48 PRELIMINARY; PRT; 469 AA.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hepatitis A virus cellular receptor 1 short form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043446; AAC39771.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 469 AA; 50561 MW; 27E386F47948F528 CRC64;

Query Match 77.5%; Score 335; DB 2; Length 469;
Best Local Similarity 80.0%; Pred. No. 7.9e-24;
Matches 64; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 2 PLPRNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSTTDGNDVTSSDGLWN 61
DB 309 PLPMQDHEPVATSPSSAQPAETHPTVLLGATRTQPTSSPLYSTTDGSDTVTSSDGLWN 368

QY 62 NNQTLFLEHSLLTANTTKG 81
DB 369 NNQTLSPHSPQMVNTTEG 388

RESULT 6
Q7JJ47
ID AC Q7JJ47 PRELIMINARY; PRT; 473 AA.
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hepatitis A virus cellular receptor 1 short form.
GN Name=HAVcr-1;
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98285791; PubMed=9621093;
RA Feigelsstock D., Thompson P., Mattoo P., Kaplan G.G.;
RT "Polymorphisms of the hepatitis A virus cellular receptor 1 in African
RT green monkey kidney cells result in antigenic variants that do not
RT react with protective monoclonal antibody 190/4.";
RL J. Virol. 72:6218-6222(1998).
DR EMBL; AF043447; AAC39772.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR SMART; SM00409; IG; 1.
DR PROSITE; PS00835; IG LIKE; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Receptor.
SQ SEQUENCE 473 AA; 51088 MW; 472DF987CA220524 CRC64;

Query Match 77.5%; Score 335; DB 2; Length 474;
Best Local Similarity 80.0%; Pred. No. 8e-24;
Matches 64; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 2 PLPRNHPEVATSPSPQPAETHPTTLOGAIRREPTSSPLYSTTDGNDVTSSDGLWN 61
DB 314 PLPMQDHEPVATSPSSAQPAETHPTVLLGATRTQPTSSPLYSTTDGSDTVTSSDGLWN 373

QY 62 NNQTLFLEHSLLTANTTKG 81
DB 374 NNQTLSPHSPQMVNTTEG 393

RESULT 8
O46598

```

ID	O46598	PRELIMINARY;	PRT;	478 AA.
AC	O46598;			
DC	01-JUN-1998 (TrEMBLrel. 06, Created)			
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Hepatitis A virus cellular receptor 1 long form.			
GN	Name=HAVcr1;			
OS	Cercopithecus aethiops (Green monkey) (Grivet).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;			
OC	Cercopitheciinae; Cercopithecus.			
OX	NCBI_TaxID=9534;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98285791; PubMed=9621093;			
RC	"Feigelson D., Thompson P., Mattoo P., Kaplan G.G.;			
RT	"Polymorphisms of the hepatitis A virus cellular receptor 1 in African			
RT	green monkey kidney cells result in antigenic variants that do not			
RL	react with protective monoclonal antibody 190/4.";			
RL	J. Virol. 72:6218-6222(1998).			
DR	EMBL; AF043449; AAC39774.1; -.			
DR	GO; GO:0004872; F:receptor activity; IEA.			
DR	InterPro; IPR003599; IG.			
DR	InterPro; IPR007110; IG-like.			
DR	InterPro; IPR003006; IG_MHC.			
DR	SMART; SM00409; IG; 1.			
DR	PROSITE; PS00835; IG LIKE; 1.			
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_1.			
KW	Receptor.			
SQ	SEQUENCE 478 AA; 51501 MW; 3A919655C752FF60 CRC64;			
Query Match 77.5%; Score 335; DB 2; Length 478;				
Best Local Similarity 80.0%; Pred. No. 8.1e-24;				
Matches 64; Conservative 4; Mismatches 12; Indels 0; Gaps 0;				
QY	2 PLPQNHEPVATSPSSPOPAETHPTTLOGAIRRPTSSPLSYVTTGNDVTSSDGLWN 61			
DB	318 PLPQMOPHEPVATSPSSPOPAETHPTTLLCATRTQTSSPLSYVTTGSDVTSSDGLWN 377			
QY	62 NNQTQLFLEHSLLTANTTKG 81			
DB	378 NNQTQLSPEHSPOMVNTTEG 397			
RESULT 9				
OS4947	PRELIMINARY;	PRT;	307 AA.	
ID	O54947;			
AC	O54947;			
DC	01-JUN-1998 (TrEMBLrel. 06, Created)			
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Kidney injury molecule-1 (KIM-1 protein).			
GN	Name=KIM-1;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Kidney;			
RC	MEDLINE=98129827; PubMed=9461608; DOI=10.1074/jbc.273.7.4135;			
RA	Ichimura T., Bonventre J.V., Bailly V., Wei H., Hession C.A.,			
RA	Cate R.L., Sanicola M;			
RT	"Kidney injury molecule-1 (KIM-1), a putative epithelial cell adhesion			
RT	molecule containing a novel immunoglobulin domain, is up-regulated in			
RT	renal cells after injury.";			
RN	J. Biol. Chem. 273:4135-4142(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Prostate;			
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Bhat N.K.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S. Carninci P., Prange C.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S. Carninci P., Prange C.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S. Carninci P., Prange C.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S. Carninci P., Prange C.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heise F.,		</	

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Vallalao D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RN SEQUENCE FROM N.A.
RC TISSUE=Colon;
RC Strausberg R.;
RA Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK025827; BAB15250.1; -;
DR ENBL; BC022217; AAH22217.1; -;
DR ENBL; AK025033; BAB15054.1; -;
KW Hypothetical protein.
SQ SEQUENCE 153 AA; 16678 MW; 11CA3A27EDCE711F CRC64;
Query Match 20.4%; Score 88; DB 2; Length 153;
Best Local Similarity 41.7%; Pred. No. 0.91;
Matches 20; Conservative 6; Mismatches 10; Indels 12; Gaps 2;
QY 3 LPRQNHVPATSSSPQAPAE-----THPTTL-----QGAIRREPTS 38
DB 46 LPBQHPFVGTAPGPNVPSSQGGRTHTPSLIRIWARRAQOGLRLPTS 93
RESULT 11
Q7PMD5 ID Q7PMD5 PRELIMINARY; PRT; 3150 AA.
AC Q7PMD5;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000004655 (Fragment).
GN Name=ENSANGG00000003651;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAA01008980; EAA13969.2; -;
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0008151; P:cell growth and/or maintenance; IEA.
DR InterPro; IPR006770; OGF_recept.
DR Pfam; PF04680; OGF_r_III; 80.
FT NON_TER 1
FT NON_TER 3150 3150
SQ SEQUENCE 3150 AA; 322879 MW; 3C7B3D441C8C839 CRC64;
Query Match 19.9%; Score 86; DB 2; Length 3150;
Best Local Similarity 34.4%; Pred. No. 57;

Matches 22; Conservative 9; Mismatches 33; Indels 0; Gaps 0;
QY 2 PLPRQNHVPATSSSPQAPAEHTTTTQGAIRREPTSSSPLYSYTTDGTGNDVTSSDGLWN 61
DB 1293 PESTTPGIRTSPTPTSTESDITMSSASTPEPTTPDTRTTRTPRTSTESTDTTMS 1352
QY 62 NNOT 65
DB 1353 SAST 1356
RESULT 12
Q75UG6 ID Q75UG6 PRELIMINARY; PRT; 642 AA.
AC Q75UG6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similar to Leishmania major. Ppg3.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum";
RL Nature 418:79-85(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR ENBL; AC116979; AAS38767.1; -;
SQ SEQUENCE 642 AA; 66631 MW; C4837BCCEFA4B962A CRC64;
Query Match 19.2%; Score 83; DB 2; Length 642;
Best Local Similarity 29.5%; Pred. No. 16;
Matches 26; Conservative 16; Mismatches 32; Indels 14; Gaps 4;
QY 2 PLPRQNHVPATSSSPQAPAEHTTTTQGAIRREPT-----SSPLYSYTTDGTGNDVTES 55
DB 461 PPTNSSSTPTPNSSTPTP--TNSSTPTPNSDPSASSFNSSSFNSTSSNSITSPS 518
QY 56 SDGLMNNNOTQLF-----LEHSLTLTANT 79
DB 519 SSSL--NNSTSSFNSTSLNSSTSSNST 544
RESULT 13
Q8SSZ4 ID Q8SSZ4 PRELIMINARY; PRT; 684 AA.
AC Q8SSZ4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum";
RL Nature 418:79-85(2002).
RN [2]

```
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116330; AAL96724.2; -.
DR HSSP; Q13952; 1N1J.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR003958; CBFA_NFYB domain.
DR InterPro; IPR009072; Histone-fold.
DR InterPro; IPR007124; Hist-TAF.
DR Pfam; PF08080; CBFD_NFYB_HMF; 1.
DR Hypothetical protein.
SQ SEQUENCE 684 AA; 78068 MW; C47DD981569820AF CRC64;
    Query Match
    Best Local Similarity 29.8%; Pred.No.19; Length 684;
    Matches 25; Conservative 10; Mismatches 20; Indels 29; Gaps 4;
QY 2 PLPQ-----NHEPVATSPSS--PQPAETHPTTLQGAIRREPTSSPLSYTTDGN 49
Db 183 PIPQPSLNDSSNGTGPALSSPSSTTPTTTPHPTT-----PTSTPNQRFQSNGS 234
QY 50 DVTSSSDGLMNNQTLFLEHSL 73
Db 235 SSF-----QNQLQNHLENKL 249
RESULT 14
Q22286 PRELIMINARY; PRT; 965 AA.
AC Q22286;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein T07C12.8.
GN ORFNames=T07C12.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
    investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurray A.A.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 273976; CAA98287.2; -.
DR PIR; T24643; T24643.
DR WormBase; WBGene00011572; T07C12.8.
DR WormPep; T07C12.8; CE28954.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000998; MAM.
DR Pfam; PF00629; MAM; 1.
DR SMART; SM00137; MAM; 1.
DR PROSITE; PS50060; MAM_2; 2.
DR Hypothetical protein.
SQ SEQUENCE 965 AA; 106620 MW; D7E2E8703D897ABA CRC64;
    Query Match
    Best Local Similarity 18.4%; Score 79.5; DB 2; Length 965;
    Matches 23; Conservative 11; Mismatches 31; Indels 23; Gaps 3;
QY 4 PRQNEPVATSSSPQPAETHPTTLQGAIRREPTSSPLSYTTDGNVTSSD----- 57
Db 498 PSNILPGGTLPSPNAPID-----VRRSEFVSS-HFYNTNQTTLTNDQBNQKTL 547
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC116330; AAL96724.2; -.
DR HSSP; Q13952; 1N1J.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR003958; CBFA_NFYB domain.
DR InterPro; IPR009072; Histone-fold.
DR InterPro; IPR007124; Hist-TAF.
DR Pfam; PF08080; CBFD_NFYB_HMF; 1.
DR Hypothetical protein.
SQ SEQUENCE 684 AA; 78068 MW; C47DD981569820AF CRC64;
    Query Match
    Best Local Similarity 29.8%; Pred.No.19; Length 684;
    Matches 25; Conservative 10; Mismatches 20; Indels 29; Gaps 4;
QY 2 PLPQ-----NHEPVATSPSS--PQPAETHPTTLQGAIRREPTSSPLSYTTDGN 49
Db 183 PIPQPSLNDSSNGTGPALSSPSSTTPTTTPHPTT-----PTSTPNQRFQSNGS 234
QY 50 DVTSSSDGLMNNQTLFLEHSL 73
Db 235 SSF-----QNQLQNHLENKL 249
RESULT 15
Q9FFW5 PRELIMINARY; PRT; 681 AA.
ID Q9FFW5;
AC Q9FFW5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Similarity to protein kinase (AT5G38560/MBB18_10).
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroids ii; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RA "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
    features of the 1.6 Mb regions covered by twenty physically assigned
    P1 clones.";
RL DNA Res. 4:215-230(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinzaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,
RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Banh J., Bowser L.,
RA Carninci P., Chang E., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Onodera C.S.,
RA Palm C.J., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
RA Tang C.C., Toriumi M., Wu H.C., Yamada K., Yamamura Y., Yu G., Yu S.,
RA Shinzaki K., Davis R.W., Theologis A., Ecker J.R.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB005231; BAB10146.1; -.
DR EMBL; AY075681; AAL77688.1; -.
DR EMBL; AF424623; AAL11616.1; -.
DR EMBL; AY113039; BAM47347.1; -.
DR HSSP; F36897; IIA5.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005199; F:structural constituent of cell wall; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase-like.
DR InterPro; IPR003882; Pistil extensin.
DR InterPro; IPR000719; Prot_kinase.
```



```

DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR PRINTS; PR01218; PSTLEXTENSIN.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 681 AA; 72389 MW; F64DA1E470E73F9 CRC64;

Query Match      18.3%; Score 79; DB 2; Length 681;
Best Local Similarity 29.8%; Pred. No. 41;
Matches 17; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

QY      2 PLPRQNHVPATSPSPQPAETHPTTLQGAIRREPTSSPLYSTTGNDTVTSSDG 58
Db      173 PPPATSGAPSSNPTDPSTLAPPPTPLPVVPREKPIAKPTGPASNNGNNTLPSSSPG 229

Search completed: June 29, 2005, 09:01:42
Job time : 34.4637 secs

```

THIS PAGE BLANK (USPTO)